

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: DARLINA DUFFY Examiner #: 72448 Date: 2/6/03
 Art Unit: 1645 Phone Number 305-7555 Serial Number: 091833,799
 Mail Box and Bldg/Room Location: 8E12 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: 6/8/90

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

091833,799 has priority to 071534,096 filed 6/8/90 and does not have to comply with the sequence rules.

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07-703-308-4498
jan.delaval@uspto.gov

Please hand enter polypeptide and coding NA structure set forth in Fig 13 attached.

Please search ① NA + polypeptide and oligomers of each.
 ② AA run against NA data bases
 ③ interference search

Please print out top 75 hits in each category

Thanks

2112/03 - 01 - 247,533 - 247,536
 63 - 30,49 - 11,791

2116/03 - 01 - 248,723

Pat Duffy

2112/03 - 01 - 458-478
 2112/03 - 01 - 416-436
 06 - 248-304 - 7 8769
 03 - 117-137

2112/03 - 01 - 327-347
 03 - 50-100
 2112/03 - 01 -
 08 - 04-84

STAFF USE ONLY

Searcher: 4498

Searcher Phone #: _____

Searcher Location: 211103

Date Searcher Picked Up: 211403

Date Completed: _____

Searcher Prep & Review Time: _____

Clerical Prep Time: 45

Online Time: 120

Type of Search

NA Sequence (#) ☒

AA Sequence (#) _____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN: _____

Dialog _____

Questel/Orbit _____

Dr. Link _____

Lexis/Nexis _____

Sequence Systems: ☒ _____

WWW/Internet _____

Other (specify) _____

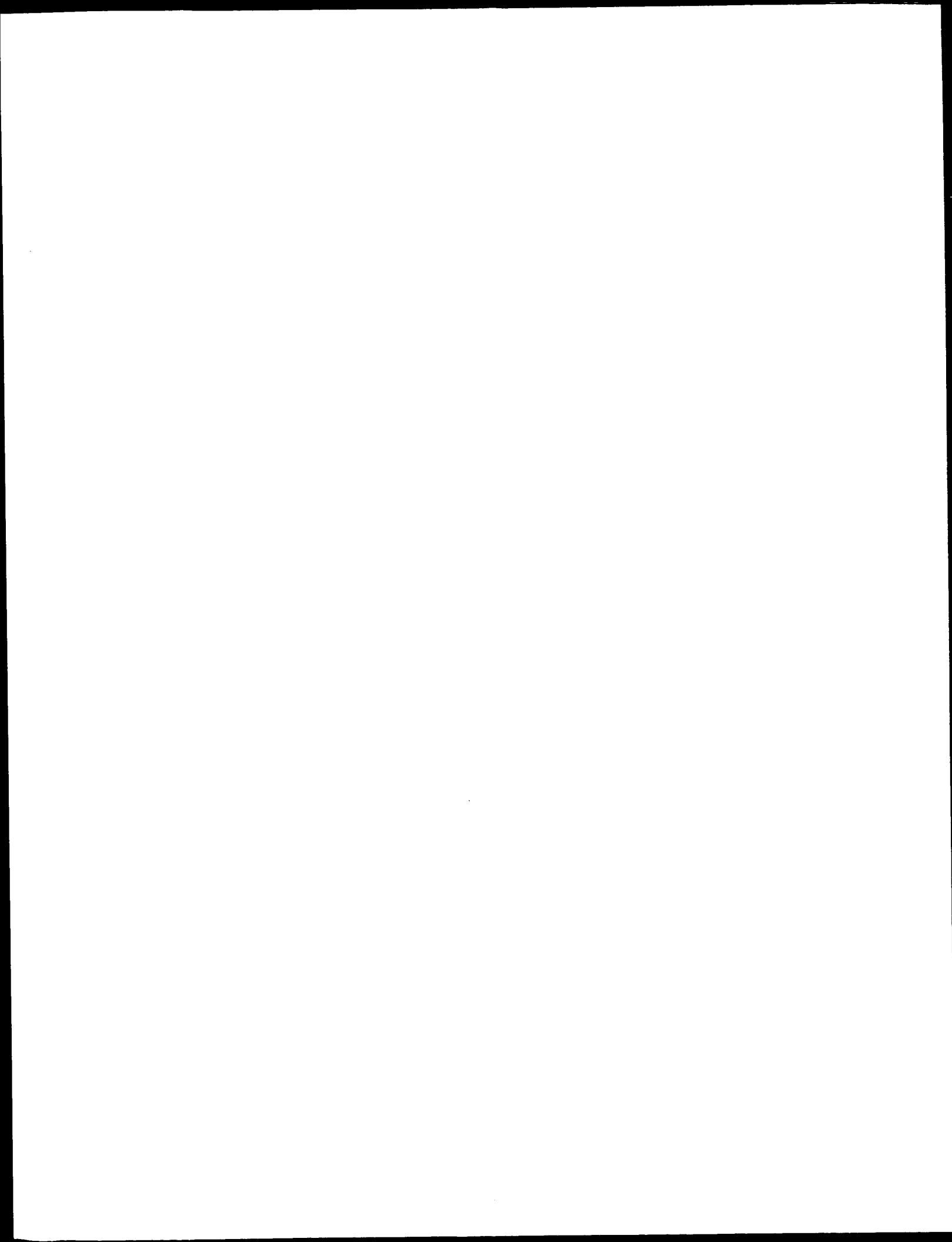


Fig. 13.

AlaGlnGluProValLysGlyProValSerThr

1

ELI1

AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAGTTAAAGGTCTTGCTACT
GCTCGAGCCATGGTATGGACGTATACGAGTTCTTGGTCAATTTCCAGGACACAGATGA

LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArg

63
 AAGCCAGGTTCTTGTCCTATTATCTTGATTTCGTTGCGCTATGTTAAACCCACCTAACCGT
 TTCGGTCCAAGAACAGGATAATAGAACTAAGCAACGCGATACAATTTGGGTGGATTGGCA
 ELI2

CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet

123
 TGT TTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTGCGGTATG
 ACAAACTTCCTGTGACTAACAGGTCCATAGTTTTTCACGACACTTCCAAGGACGCCATAC
 ELI5
 ELI4

AlaCysPheValProGlnEndEnd

183 GCTTGTTTCGTTCCACAATAATAG

CGAACAAGCAAGGTGTTATTATCCTAG 210

ELI6



12-1-77

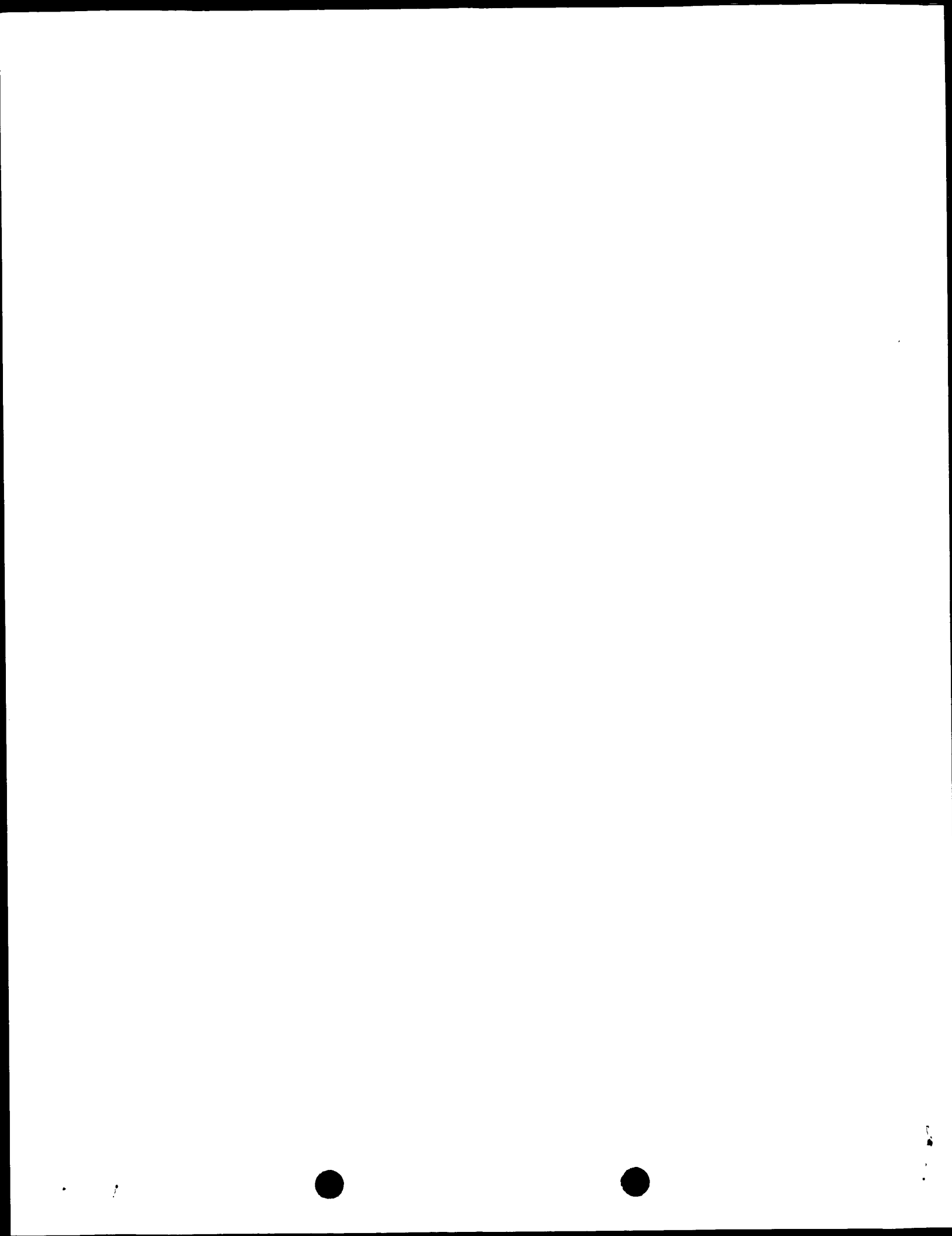
Sun Feb 16 10:28:26 2003

US09833799.pep

Page 1

; Entered [jdelaval 11-Feb-03 8:13]
09-833799-13B
agepvkpvatkp9scpiilircamlnppnrciktdcp9gikkcccgscgmacfvpq1

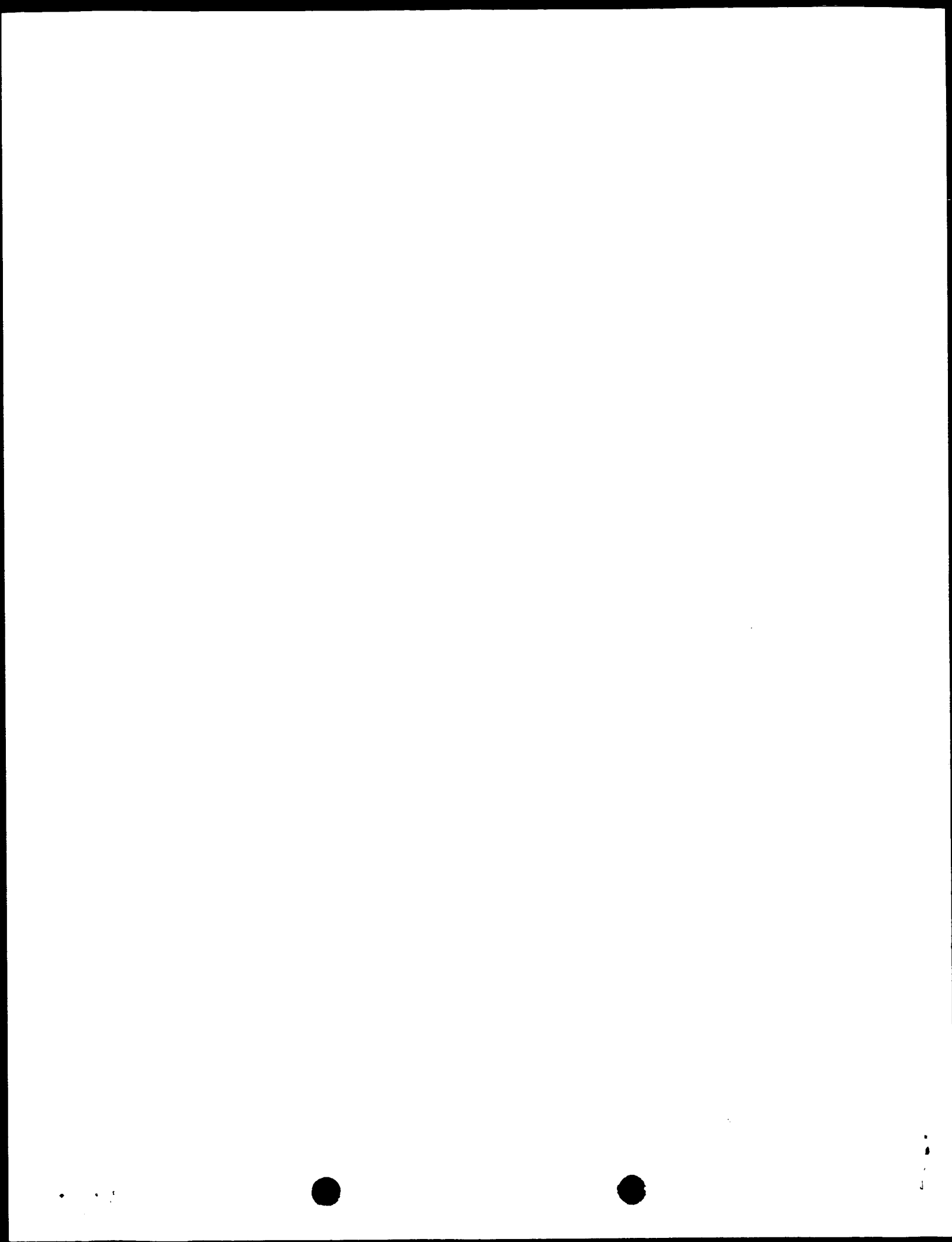
Jar Delaval
Reference Librarian
Technology & Chemical Library
C/M 1E07 - 703-308-4498
jdelaval@uspto.gov



: Entered [jdelaval 11-Feb-03 9:38]
09-833799-13A
AATTCGAGCTCGGTACCATACCTGCAATATGCTCAAGAACCAAGTTAAAGGTCCTGTGTCTACTAAGCCAGG
TTCTTGTCCTATTATCTTGATTCTGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGAT
TGTCAGGTATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCTTGTTTCGTTCACAAATATATAG1

: Entered [jdelaval 11-Feb-03 9:55]
09-833799-13C
gctcgagccatgggtatggagcgtatcacgagttcttgggtcaatttccaggacacacagatgattcggtccaaga
acaggataatggaactaagcaacgcgatacaaatttgggtggattggcaaaaacttccctgtgactaacag
gtccatagtttttccagcacacttccaaggacgcctaccggaacaagcaaggtgtattatccttag1

Jean Delaval
Science Librarian
Biology & Chemical Library
E07 - 703-308-4498
jdelaval@uspto.gov



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 08:35:56 ; Search time 10 Seconds
(without alignments)
236.415 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 agepvgpvtkpqscpiil.....cpgikkccgscgmactvpq 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	117	1	ELAF_HUMAN
2	221.5	67.7	167	1	ELAF_PIG
3	179.5	54.9	187	1	SPAI_PIG
4	165	50.5	115	1	ALK1_PIG
5	161	49.2	131	1	ALK1_MOUSE
6	160	48.9	144	1	WAP3_PIG
7	158	48.3	55	1	CALU_CAVPO
8	143	43.7	132	1	ALK1_HUMAN
9	101.5	31.0	74	1	WDMN_RAT
10	100.5	30.7	123	1	WFD2_RABIT
11	100.5	30.7	191	1	WAP_MACEU
12	100.5	30.7	676	1	KALM_CHICK
13	98	30.0	137	1	WAP_RAT
14	96.5	29.5	174	1	WDMN_MOUSE
15	95.5	29.2	132	1	WAP_PIG
16	93	28.4	124	1	WFD2_HUMAN
17	92.5	28.3	680	1	KALM_HUMAN
18	90.5	27.7	117	1	WAP_CAMDR
19	88	26.9	124	1	WFD1_CANFA
20	88	26.9	134	1	EPPI_MOUSE
21	85	26.0	133	1	EPPI_HUMAN
22	85	26.0	133	1	EPPI_MACMU
23	81.5	24.9	212	1	WFD1_RAT
24	80.5	24.6	211	1	WFD1_MOUSE
25	79.5	24.3	134	1	WAP_MOUSE
26	79.5	24.3	220	1	WFD1_HUMAN
27	78	23.9	127	1	WAP_RABIT
28	74.5	22.8	110	1	IBP_CARCR
29	67	20.5	34	1	ITR1_MOMCO
30	66	20.2	2911	1	FBN2_HUMAN
31	65	19.9	1700	1	BAR3_CHITE
32	64	19.6	30	1	ITR3_MOMCO
33	64	19.6	34	1	ITR2_MOMCO

34	63	19.3	2907	1	FBN2_MOUSE
35	62.5	19.1	202	1	ADEN_ADEB7
36	62.5	19.1	1964	1	NTC4_MOUSE
37	62	19.0	71	1	MT2A_MYTE
38	62	19.0	71	1	MT2B_MYTE
39	61.5	18.8	201	1	ADEN_ADE07
40	61.5	18.8	1363	1	ILPR_BRALA
41	61	18.7	129	1	TFP2_HUMAN
42	61	18.7	1246	1	YMW2_CABEL
43	60	18.3	30	1	ITR1_MOMCH
44	60	18.3	71	1	MT21_MYTE
45	60	18.3	71	1	MT22_MYTE
46	60	18.3	5376	1	ZAN_MOUSE
47	59.5	18.2	75	1	MT_EISFO
48	59.5	18.2	621	1	YC92_CABEL
49	59.5	18.2	2003	1	NTC4_HUMAN
50	59	18.0	2871	1	FBN1_MOUSE
51	59	18.0	4289	1	TENX_HUMAN
52	58	17.7	129	1	TFP2_MOUSE
53	58	17.7	694	1	AGSA_CRYPV
54	58	17.7	2214	1	SORL_HUMAN
55	58	17.7	2813	1	VMP_CANFA
56	57.5	17.6	74	1	MT_CRAVI
57	57.5	17.6	204	1	ADEN_ADEF1
58	57.5	17.6	4660	1	LRP2_RAT
59	57	17.4	954	1	M3KA_HUMAN
60	56.5	17.3	28	1	IEL1_MOMCH
61	56.5	17.3	61	1	MT1B_HUMAN
62	56.5	17.3	185	1	RL25_CHLTR
63	56.5	17.3	326	1	YAW5_SCHPO
64	56	17.1	70	1	CX2X_CONBE
65	56	17.1	241	1	TR18_HUMAN
66	56	17.1	1342	1	ERB3_HUMAN
67	56	17.1	1680	1	FUR2_DROME
68	56	17.1	2871	1	FBN1_HUMAN
69	55.5	17.0	78	1	HSTN_VIBCH
70	55.5	17.0	78	1	HSTN_VIBCH
71	55.5	17.0	129	1	TFP2_RAT
72	55.5	17.0	287	1	TMEP_HUMAN
73	55.5	17.0	500	1	U7I3_HUMAN
74	55.5	17.0	518	1	VG47_HSV11
75	55	16.8	32	1	ITR4_CUCWA
76	55	16.8	46	1	EN2A_HORSE
77	55	16.8	855	1	ST14_HUMAN
78	55	16.8	1207	1	EGF_HUMAN
79	54.5	16.7	28	1	ITR3_MOMCH
80	54.5	16.7	329	1	APPF_BACSU
81	54.5	16.7	462	1	RD21_ARATH
82	54.5	16.7	559	1	FZD5_XENLA
83	54.5	16.7	684	1	ACS2_KLJLA
84	54	16.5	65	1	MTB_STRPU
85	54	16.5	127	1	TFP2_PIG
86	54	16.5	327	1	Y627_CHLTR
87	54	16.5	625	1	XYNA_PIRSP
88	54	16.5	769	1	ITB2_BOVIN
89	54	16.5	1339	1	ERB3_RAT
90	54	16.5	2703	1	NOTC_DROME
91	54	16.5	2871	1	FBN1_BOVIN
92	54	16.5	2871	1	FBN1_PIG
93	54	16.5	3767	1	MUA3_CABEL
94	54	16.5	4753	1	LRP_CABEL
95	53.5	16.4	157	1	NEU1_ONCMA
96	53.5	16.4	328	1	CI17_GIALA
97	53.5	16.4	346	1	CYSL_LYCES
98	53.5	16.4	451	1	DADR_XENLA
99	53.5	16.4	478	1	DISR_AGRKH
100	53.5	16.4	604	1	CFAI_RAT

ALIGNMENTS

ELAF_HUMAN
ID ELAF_HUMAN STANDARD; PRT; 117 AA.
AC P19957;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elafin precursor (Elastase-specific inhibitor) (ESI) (Skin-derived
DN antileukoprotease) (SKALP).
GN P13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92287100; PubMed=1339270;
RA Saheki T., Ito F., Hagiwara H., Saito Y., Kuroki J., Tachibana S.,
HIROSE S.;
RT "Primary structure of the human elafin precursor proelafin deduced
RT from the nucleotide sequence of its gene and the presence of unique
RT repetitive sequences in the prosegment.";
RL Biochem. Biophys. Res. Commun. 185:240-245 (1992).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-48 AND 85-100.
RC TISSUE=Keratinocytes;
RX MEDLINE=93280175; PubMed=7685029;
RA Molhuizen H.O.F., Alkemade H.A.C., Zeeuwen P.L.J.M.,
de Jongh G.J., Wieringa B., Schalkwijk J.;
RT "SKALP/elafin: an elastase inhibitor from cultured human
RT keratinocytes. Purification, cDNA sequence, and evidence for
RT transglutaminase cross-linking.";
RL J. Biol. Chem. 268:12028-12032 (1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93236929; PubMed=8476637;
RA Sallenave J.M., Silva A.;
RT "Characterization and gene sequence of the precursor of elafin, an
RT elastase-specific inhibitor in bronchial secretions.";
RL Am. J. Respir. Cell Mol. Biol. 8:439-445 (1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
Coulson A., Coville G.J., Deadman R., Dhama P.D., Dunn M.,
Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
Lehvasaaho M.H., Loversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
Marsh V.L., Martin S.L., McConnachie L.J., McElay K., McMurray A.A.,
Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871 (2001).
RN [5]
RP SEQUENCE OF 61-117.
RC TISSUE=Horny layers;

RX MEDLINE=90368643; PubMed=2394696;
RA Wiedow O., Schroeder J.-M., Gregory H., Young J.A., Christophers E.;
RT "Elafin: an elastase-specific inhibitor of human skin. Purification,
RT characterization, and complete amino acid sequence.";
RL J. Biol. Chem. 265:14791-14795 (1990).
RN [6]
RP ERRATUM.
RA Wiedow O., Schroeder J.-M., Gregory H., Young J.A., Christophers E.;
RL J. Biol. Chem. 266:3356-3356 (1991).
RN [7]
RP SEQUENCE OF 70-94.
RC TISSUE=Sputum;
RX MEDLINE=91248412; PubMed=2039600;
RA Sallenave J.-M., Ryle A.P.;
RT "Purification and characterization of elastase-specific inhibitor.
RT Sequence homology with mucus proteinase inhibitor.";
RL Biol. Chem. Hoppe-Seyler 372:13-21 (1991).
RN [8]
RP SEQUENCE OF 61-78.
RC TISSUE=Sputum;
RX MEDLINE=92162196; PubMed=1536690;
RA Sallenave J.-M., Marsden M.D., Ryle A.P.;
RT "Isolation of elafin and elastase-specific inhibitor (ESI) from
RT bronchial secretions. Evidence of sequence homology and immunological
RT cross-reactivity.";
RL Biol. Chem. Hoppe-Seyler 373:27-33 (1992).
RN [9]
RP SEQUENCE OF 85-100.
RC TISSUE=Keratinocytes;
RX MEDLINE=91159523; PubMed=2001428;
RA Schalkwijk J., de Roo C., de Jongh G.J.;
RT "Skin-derived antileukoprotease (SKALP), an elastase inhibitor from
RT human keratinocytes. Purification and biochemical properties.";
RL Biochim. Biophys. Acta 1096:148-154 (1991).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96387196; PubMed=8794736;
RA Tsunemi M., Matsura Y., Sakakibara S., Katsube Y.;
RT "Crystal structure of an elastase-specific inhibitor elafin complexed
RT with porcine pancreatic elastase determined at 1.9-A resolution.";
RL Biochemistry 35:11570-11576 (1996).
RN [11]
RP STRUCTURE BY NMR.
RX MEDLINE=97315116; PubMed=9171290;
RA Francart C., Dauchez M., Alix A.J.P., Lippens G.;
RT "Solution structure of R-elafin, a specific inhibitor of elastase.";
RL J. Mol. Biol. 268:666-677 (1997).
CC -!- FUNCTION: NEUTROPHIL AND PANCREATIC ELASTASE-SPECIFIC INHIBITOR OF
CC SKIN. IT MAY PREVENT ELASTASE-MEDIATED TISSUE PROTEOLYSIS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: CONSISTS OF TWO DOMAINS: THE TRANSGLUTAMINASE SUBSTRATE
CC DOMAIN (CEMENTOIN MOIETY) AND THE ELASTASE INHIBITOR DOMAIN. THE
CC TRANSGLUTAMINASE SUBSTRATE DOMAIN SERVES AS AN ANCHOR TO LOCALIZE
CC ELAFIN COVALENTLY TO SPECIFIC SITES ON EXTRACELLULAR MATRIX
CC PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z18538; CAA79223.1; -
CC EMBL: D13156; BAA02441.1; -
CC EMBL: S58717; AAB26371.1; -
CC EMBL: L10343; AAA36483.1; -
CC EMBL: AL049767; CAB53524.1; -
CC PIR: JH0614; JH0614.
CC PIR: S13607; S13607.
CC PIR: A46749; A46749.


```

DR PDB; 1FLE; 11-JAN-97.
DR PDB; 2REL; 07-JUL-97.
DR Genew; HGNC:8947; P13.
DR MIM; 182257; -.
DR InterPro; IPR002098; SVP 1.
DR InterPro; IPR002221; WAP-.
DR Pfam; PF000095; wap; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00313; SVP 1; 2.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
DR Serine protease inhibitor; Signal; Repeat; 3D-structure.
FT SIGNAL 1 22
FT PROPEP 23 60
FT CHAIN 61 117
FT DOMAIN 29 72
FT REPEAT 29 72
FT REPEAT 55 72
FT REPEAT 72 117
FT DOMAIN 76 105
FT DISULFID 83 109
FT DISULFID 92 104
FT DISULFID 98 113
FT CONFLICT 91 92
FT CONFLICT 92 92
FT CONFLICT 98 98
SQ SEQUENCE 117 AA; 12270 MW; A51D46257E5B03EF CRC64;

Query Match 100.0%; Score 327; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.3e-29;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQPVGKPGVSTKPGSPILLIRCAMLNPNRCLKDTDCPGIKKCCGSGCMACFPVQ 57
Db 61 AQPVGKPGVSTKPGSPILLIRCAMLNPNRCLKDTDCPGIKKCCGSGCMACFPVQ 117

RESULT 2
ELAF_PIG STANDARD; PRT; 167 AA.
AC Q29125;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elafin precursor (WAP-1 protein).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215132; PubMed=8636131;
RA Tamechika I., Itakura M., Saruta Y., Furukawa M., Kato A.,
RA Tachibana S., Hirose S.;
RT "Accelerated evolution in inhibitor domains of porcine elafin family members."
RL J. Biol. Chem. 271:7012-7018 (1996).
CC -!- FUNCTION: NEUTROPHIL AND PANCREATIC ELASTASE-SPECIFIC INHIBITOR OF SKIN. IT MAY PREVENT ELASTASE-MEDIATED TISSUE PROTEOLYSIS (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: TRACHEA AND LARGE INTESTINE.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50319; BAA08854.1; -.
DR HSSP; P19957; 2REL.

```

```

DR InterPro; IPR002221; WAP.
DR Pfam; PF000095; wap; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
DR Serine protease inhibitor; Signal; Repeat.
FT SIGNAL 1 21
FT PROPEP 22 70
FT CHAIN 71 167
FT DOMAIN 44 115
FT REPEAT 44 49
FT REPEAT 50 55
FT REPEAT 56 61
FT REPEAT 62 67
FT REPEAT 68 73
FT REPEAT 74 79
FT REPEAT 80 85
FT REPEAT 86 91
FT REPEAT 92 97
FT REPEAT 98 103
FT REPEAT 104 109
FT REPEAT 110 115
FT DOMAIN 78 126
FT REPEAT 78 101
FT REPEAT 102 126
FT DOMAIN 122 167
FT DISULFID 126 155
FT DISULFID 133 159
FT DISULFID 142 154
FT DISULFID 148 163
SQ SEQUENCE 167 AA; 17923 MW; 25932EA1A459D9CA CRC64;

Query Match 67.7%; Score 221.5; DB 1; Length 167;
Best Local Similarity 61.9%; Pred. No. 1.7e-17;
Matches 33; Conservative 5; Mismatches 12; Indels 7; Gaps 1;

QY 2 QBPVKGP-----VSTKPGSPILLIRCAMLNPNRCLKDTDCPGIKKCCGSGCMACF 54
Db 105 QBPVKGAQPAIKRLILLTKPGSPILLIRCAMLNPNRCLSDAQCPLKKCCGFCGKACM 164

QY 55 VPQ 57
Db 165 DPK 167

RESULT 3
SPAI_PIG STANDARD; PRT; 187 AA.
AC P16225;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sodium/potassium ATPase inhibitor SPAI-2 precursor (WAP-2 protein).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-58.
RC TISSUE=Duodenum;
RX MEDLINE=95403443; PubMed=7673229;
RA Kuroki J., Hosoya T., Itakura M., Hirose S., Tamechika I.,
RA Yoshimoto T., Ghoneim M.A., Nara K., Kato A., Suzuki Y., Furukawa M.,
RA Tachibana S.;
RT "Cloning, characterization, and tissue distribution of porcine SPAI, a protein with a transglutaminase substrate domain and the WAP motif."
RL J. Biol. Chem. 270:22428-22433 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215132; PubMed=8636131;
RA Tamechika I., Itakura M., Saruta Y., Furukawa M., Kato A.,
RA Tachibana S., Hirose S.;

```

"Accelerated evolution in inhibitor domains of porcine elafin family members.";
J. Biol. Chem. 271:7012-7018(1996).

[3]

SEQUENCE OF 127-187.

TISSUE=Duodenum;

MEDLINE=90026425; PubMed=2553020;

Araki K., Kuroki J., Ito O., Kuwada M., Tachibana S.;

"Novel peptide inhibitor (SPAI) of Na⁺, K⁺-ATPase from porcine

intestine.";

Biochem. Biophys. Res. Commun. 164:496-502(1989).

[4]

DISULFIDE BONDS.

MEDLINE=91025070; PubMed=2171523;

Araki K., Kuwada M., Ito O., Kuroki J., Tachibana S.;

"Four disulfide bonds' allocation of Na⁺, K⁺-ATPase inhibitor

(SPAI).";

Biochem. Biophys. Res. Commun. 172:42-46(1990).

-!- FUNCTION: INHIBITS Na⁺, K⁺ ATPASE BY THE COMPETITIVE MODE

AGAINST Na⁺.

-!- TISSUE SPECIFICITY: SMALL INTESTINE > LARGE INTESTINE. THE PLASMA

CONTAINS THE PRO-SPAI FORM CIRCULATING.

-!- DOMAIN: THE REPETITIVE DOMAIN OF PRO-SPAI SERVES AS A SUBSTRATE

FOR TRANSGLUTAMINASE.

-!- PTM: THE SHORT FORM (AA 127-187) MAY BE AN ARTIFACT DUE TO THE

STRONGLY ACIDIC CONDITIONS OF THE DUODENUM. THE PRO-SPAI FORM MAY

BE THE NATIVE FORM.

-!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D17756; BAA04603.1; -.

EMBL; D17754; BAA04603.1; JOINED.

EMBL; D17755; BAA04603.1; JOINED.

EMBL; D83667; BAA12037.1; -.

EMBL; D50320; BAA08855.1; -.

PIR; A33429; A33429.

PIR; B33429; B33429.

PIR; C33429; C33429.

HSP; P19957; 2REL.

InterPro; IPR002098; SVP 1.

InterPro; IPR002221; WAP.

PRINTS; PR00095; wap; 1.

PRINTS; PR00003; 4DISULPHCORE.

SMART; SM00217; WAP; 1.

PROSITE; PS00313; SVP 1; 2.

PROSITE; PS00317; 4_DISULFIDE_CORE; 1.

Signal; Repeat 1

Signal; Repeat 21

PROPEP 22 126

CHAIN 127 187

REPEAT 62 85

REPEAT 34 129

DOMAIN 34 39

REPEAT 34 39

REPEAT 40 45

REPEAT 46 51

REPEAT 58 63

REPEAT 64 69

REPEAT 70 75

REPEAT 76 81

REPEAT 82 87

REPEAT 88 93

REPEAT 100 105

REPEAT 106 111

REPEAT 112 117

REPEAT 118 123

REPEAT 124 129

SODIUM/POTASSIUM ATPASE INHIBITOR SPAI-2.

SVP-1 CLOTTING 1.

14 X 6 AA APPROXIMATE TANDEM REPEATS.

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

11.

12.

13.

14.

FT DOMAIN 142 187 WAP.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 146 175
FT DISULFID 153 179
FT DISULFID 162 174
FT DISULFID 168 183
FT VARIANT 127 138 MISSING (IN SPAI-1).
FT VARIANT 148 148 R -> G (IN SPAI-3).
FT VARIANT 156 156 S -> G (IN SPAI-3).
SQ SEQUENCE 187 AA; 20471 MW; 12F5BD0813AF5E27 CRC64;
Query Match 54.9%; Score 179.5; DB 1; Length 187;
Best Local Similarity 49.2%; Pred. No. 6.7e-13;
Matches 31; Conservative 8; Mismatches 17; Indels 7; Gaps 1;
QY 2 QBPVKGK-----VSTKPGSCPIILIRCAMLNPNRCLKDTDCPGIKKCCGSCGMACF 54
Db 125 QDPVKAQPAVGRFLSLRGHCFRLFCPLSNPNKRCWRDYDCGVKKCCGCGKDCDL 184
QY 55 VPQ 57
Db 185 YPK 187

RESULT 4

ALK1_PIG

ID ALK1_PIG STANDARD; PRT; 115 AA.

AC P22298;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Antileukoprotease.

GN ALP.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OC NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RX MEDLINE=91155942; PubMed=2293019;

RA Farmer S.J., Fliss A.E., Simmen R.C.M.;

"Complementary DNA cloning and regulation of expression of the

messenger RNA encoding a pregnancy-associated porcine uterine protein

related to human antileukoprotease.";

Mol. Endocrinol. 4:1095-1104(1990).

CC -!- FUNCTION: THIS MUCOSAL SECRETORY PROTEIN MAY HAVE A SPECIFIC

FUNCTION DURING FETAL DEVELOPMENT.

CC -!- TISSUE SPECIFICITY: UTERUS AND LUNG.

CC -!- DEVELOPMENTAL STAGE: MAXIMAL EXPRESSION AT MID- AND LATE

GESTATION.

CC -!- INDUCTION: BY ESTROGEN AND PROGESTERONE; IN UTERUS.

CC -!- PTM: COMPARED TO HUMAN ALP, IT SEEMS TO LACK A CLEAVABLE

SIGNAL PEPTIDE

CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M57446; AAA63446.1; -.

PIR; A36113; A36113.

HSP; P19957; 2REL.

InterPro; IPR002221; WAP.

PRINTS; PR00095; wap; 2.

PRINTS; PR00003; 4DISULPHCORE.

PRODom; PD001224; WAP; 1.

SMART; SM00217; WAP; 2.

PROSITE; PS00317; 4_DISULFIDE_CORE; 2.

KW Serine protease inhibitor; Repeat.
 FT DOMAIN 8 62 TRYPsin INHIBITORY DOMAIN.
 FT ELASTASE INHIBITORY DOMAIN.
 FT DOMAIN 14 59 WAP 1.
 FT DOMAIN 68 113 WAP 2.
 FT ACT SITE 28 28 TRYPsin INHIBITORY SITE (PROBABLE).
 FT ACT_SITE 81 82 ELASTASE OR CHYMOTRYPsin INHIBITORY SITE (PROBABLE).
 FT DISULFID 18 47 BY SIMILARITY.
 FT DISULFID 26 51 BY SIMILARITY.
 FT DISULFID 34 46 BY SIMILARITY.
 FT DISULFID 40 55 BY SIMILARITY.
 FT DISULFID 72 101 BY SIMILARITY.
 FT DISULFID 79 105 BY SIMILARITY.
 FT DISULFID 88 100 BY SIMILARITY.
 FT DISULFID 94 109 BY SIMILARITY.
 SQ SEQUENCE 115 AA; 12518 MW; C8C937E997815BCF CRC64;
 Query Match 50.5%; Score 165; DB 1; Length 115;
 Best Local Similarity 48.1%; Pred. No. 1.7e-11;
 Matches 25; Conservative 6; Mismatches 21; Indels 0; Gaps 0;
 Qy 5 VKGPVSTKPGSCPIILIRCAMLNPPNRCIKDTPCGIKKCEGSCGCMACFPV 56
 Db 61 ITNPKVKGKPGCVVYGCWMLNPPNRCIKDTPCGIKKCEGSCGCMACFPV 112
 RESULT 5
 ALKI MOUSE STANDARD; PRT; 131 AA.
 AC P97430; O09081; O09082;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antileukoprotease 1 precursor (ALP) (Secretory leukocyte protease inhibitor).
 DE SLPI.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97191310; PubMed=9039268;
 RA Jin F.-Y., Nathan C.F., Radzioch D., Ding A.;
 RT "Secretory leukocyte protease inhibitor: a macrophage product induced by and antagonistic to bacterial lipopolysaccharide.";
 RL Cell 88:417-426(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97271386; PubMed=9126337;
 RA Zlatnik R.J., Zhang J., Kashem M.A., Kohno T., Lyons D.E., Wright C.D., Rosen E., Goldberg I., Hayday A.C.;
 RT "The cloning and characterization of a murine secretory leukocyte protease inhibitor cDNA.";
 RL Biochem. Biophys. Res. Commun. 232:687-697(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
 RX MEDLINE=98011992; PubMed=9351627;
 RA Abe T., Tomimaga Y., Kikuchi T., Watanabe A., Satoh K., Watanabe Y., Nukiwa T.;
 RT "Bacterial pneumonia causes augmented expression of the secretory leukoprotease inhibitor gene in the murine lung.";
 RL Am. J. Respir. Crit. Care Med. 156:1235-1240(1997).
 CC -!- FUNCTION: ACID-STABLE PROTEINASE INHIBITOR WITH STRONG AFFINITIES FOR TRYPsin, CHYMOTRYPsin, ELASTASE, AND CATHEPSIN G. MAY PREVENT ELASTASE-MEDIATED DAMAGE TO ORAL AND POSSIBLY OTHER MUCOSAL TISSUES (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN LUNG, SPLEEN, INTESTINE AND EPIDIDYMIS WITH LOWER LEVELS IN LIVER AND SEMINAL VESICLE. NO EXPRESSION IN BRAIN, HEART, KIDNEY AND MUSCLE.
 CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U73004; AAC53047.1; -;
 CC EMBL; U88093; AAC53140.1; -;
 CC EMBL; U94341; AAC53394.1; -;
 CC HSSP; P19957; 2REL.
 CC MGD; MGI:109297; SLP1.
 CC InterPro; IPR002221; WAP.
 CC Pfam; PF00095; wap; 2.
 CC ProDom; PD001224; wap; 1.
 CC SMART; SM00217; WAP; 2.
 CC PROSITE; PS00317; 4_DISULFIDE_CORE; 2.
 KW Serine protease inhibitor; Repeat; Signal.
 FT SIGNAL 1 25 BY SIMILARITY.
 FT CHAIN 26 131 ANTILEUKOPROTEINASE 1.
 FT DOMAIN 26 84 TRYPsin INHIBITORY DOMAIN.
 FT DOMAIN 85 131 ELASTASE INHIBITORY DOMAIN.
 FT DOMAIN 32 77 WAP 1.
 FT DOMAIN 86 131 WAP 2.
 FT ACT SITE 46 46 TRYPsin INHIBITORY SITE (PROBABLE).
 FT ACT_SITE 98 99 ELASTASE OR CHYMOTRYPsin INHIBITORY SITE (PROBABLE).
 FT DISULFID 36 65 BY SIMILARITY.
 FT DISULFID 44 69 BY SIMILARITY.
 FT DISULFID 52 64 BY SIMILARITY.
 FT DISULFID 58 73 BY SIMILARITY.
 FT DISULFID 90 119 BY SIMILARITY.
 FT DISULFID 97 123 BY SIMILARITY.
 FT DISULFID 106 118 BY SIMILARITY.
 FT DISULFID 112 127 BY SIMILARITY.
 SQ SEQUENCE 131 AA; 14308 MW; A57C9E30FE711B9F CRC64;
 Query Match 49.2%; Score 161; DB 1; Length 131;
 Best Local Similarity 52.8%; Pred. No. 5.1e-11;
 Matches 28; Conservative 3; Mismatches 22; Indels 0; Gaps 0;
 Qy 4 PVKGPVSTKPGSCPIILIRCAMLNPPNRCIKDTPCGIKKCEGSCGCMACFPV 56
 Db 78 PIRKPVWKGKPGCVKVTQRCWMLNPPNRCIKDTPCGIKKCEGSCGCMACFPV 130
 RESULT 6
 WAP3_PIG STANDARD; PRT; 144 AA.
 AC Q29126;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE WAP-3 protein precursor.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215132; PubMed=8636131;
 RA Tamechika I., Itakura M., Saruta Y., Furukawa M., Kato A., Tachibana S., Hirose S.;
 RT "Accelerated evolution in inhibitor domains of porcine elafin family members.";
 RL J. Biol. Chem. 271:7012-7018(1996).
 CC -!- TISSUE SPECIFICITY: LARGE INTESTINE (RELATIVELY LOW LEVELS).
 CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.,
 "The DNA sequence and comparative analysis of human chromosome 20.",
 Nature 414:865-871(2001).
 [5]
 SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RC Strausberg R.;
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 SEQUENCE OF 26-132, AND SEQUENCE OF 26-65 FROM N.A.
 RP MEDLINE=86164996; PubMed=3485543;
 RX Seemueller U., Arnholt M., Fritz H., Wiedenmann K., Machleidt W., Heintel R., Appelhaus H., Gassen H.-G., Lottspeich F.;
 RA "The acid-stable proteinase inhibitor of human mucous secretions (HUSI-1, antileukoprotease). Complete amino acid sequence as revealed by protein and cDNA sequencing and structural homology to whey proteins and Red Sea turtle proteinase inhibitor.",
 FEBS Lett. 199:43-48(1986).
 RT [7]
 SEQUENCE OF 26-132.
 RP MEDLINE=86313644; PubMed=3462719;
 RX Thompson R.C., Ohlsson K.;
 RA "Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor, a potent inhibitor of leukocyte elastase.",
 Proc. Natl. Acad. Sci. U.S.A. 83:6692-6696(1986).
 RL [8]
 SEQUENCE OF 26-52.
 RP MEDLINE=91248412; PubMed=2039600;
 RX Sallenne J.-M., Ryle A.P.;
 RA "Purification and characterization of elastase-specific inhibitor. Sequence homology with mucus proteinase inhibitor",
 Biol. Chem. Hoppe-Seyler 372:13-21(1991).
 RL [9]
 X-RAY CRYSTALLOGRAPHY.
 RP MEDLINE=88211544; PubMed=3366116;
 RX Gruetter M.G., Fendrich G., Huber R., Bode W.;
 RA "The 2.5 A X-ray crystal structure of the acid-stable proteinase inhibitor from human mucous secretions analysed in its complex with bovine alpha-chymotrypsin.",
 EMBO J. 7:345-351(1988).
 RT CC -!- FUNCTION: ACID-STABLE PROTEINASE INHIBITOR WITH STRONG AFFINITIES FOR TRYPSIN, CHYMOTRYPSIN, ELASTASE, AND CATHEPSIN G. MAY PREVENT ELASTASE-MEDIATED DAMAGE TO ORAL AND POSSIBLY OTHER MUCOSAL TISSUES.
 CC
 CC -!- TISSUE SPECIFICITY: MUCOUS FLUIDS.
 CC
 CC -!- DISEASE: THE PATHOLOGIES OF SEVERAL CHRONIC AND ACUTE DISEASES OF THE RESPIRATORY TRACT INVOLVE AN IMBALANCE BETWEEN THE PROTEASES OF CELLS INVOLVED IN INFLAMMATORY RESPONSES & THE INHIBITORS OF THESE PROTEASES. THE INFLAMMATORY-MEDIATED RELEASE OF NEUTROPHIL ELASTASE IN THE LUNGS OF PATIENTS WHOSE LEVELS OF ACTIVE ALPHA-1-ANTITRYPSIN ARE COMPROMISED BY GENETIC BACKGROUND, CIGARETTE SMOKING, AIR POLLUTANTS, OR A COMBINATION OF ALL THREE CAN RESULT IN SEVERE LUNG DAMAGE AND A DECREASED LIFESPAN. THE RELATIVELY SMALL SIZE OF THIS PROTEIN, ITS LACK OF GLYCOSYLATION AND ITS STABILITY MAKE THIS PROTEIN A CANDIDATE FOR USE AS A THERAPEUTIC AGENT IN DISEASES MEDIATED BY LEUKOCYTE ELASTASE-ANTI-ELASTASE IMBALANCES.
 CC
 CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch.
 CC
 CC EMBL; X04470; CAA28158.1; -
 DR

RA Dear T.N., Kefford R.F.;
RT "The WDM1 gene product is a novel member of the 'four-disulphide
RL core' family of proteins."
RL Biochem. Biophys. Res. Commun. 176:247-254(1991).
CC -!- FUNCTION: INVOLVED IN THE METASTATIC POTENTIAL OF ADENOCARCINOMAS
CC -!- IN RAT. COULD HAVE PROTEINASE INHIBITING CAPACITY.
CC -!- TISSUE SPECIFICITY: SPLEEN, KIDNEY, BRAIN, LIVER, LUNG, AND CELL
CC LINE RAT-2.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13309; CAA31688.1; ALT_INIT.
DR F1R; JHO390; JHO390.
DR HSP; P19957; 2REL.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 1.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4-DISULFIDE_CORE; 1.
KW Protease inhibitor; Signal.
FT NON_TER 1
FT SIGNAL <1 23 POTENTIAL.
FT CHAIN 24 74 WDM1 PROTEIN.
FT DOMAIN 29 73 WAP.
FT SEQUENCE 74 AA; 7740 MW; 51802C70CDAP0521 CRC64;
SQ

Query Match 31.0%; Score 101.5; DB 1; Length 74;
Best Local Similarity 38.5%; Pred. No. 9.3e-05;
Matches 20; Conservative 1; Mismatches 16; Indels 15; Gaps 2;

QY 12 KPGSCPIILIRCAMLPNPNR-----CLKDTPCPIKCKCGSCGMACFVP 56
DB 29 KPGKCP-----KNPPRSIGTCVELSGDQSCFNQKCCSGCHVKSP 72

RESULT 10
WFD2_RABIT
ID WFD2_RABIT STANDARD; PRT; 123 AA.
AC Q28631;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major epididymis-specific protein E4 precursor (Epididymal protein BE-
DE 20) (WAP four-disulphide core domain protein 2).
GN WFD2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND REVISION TO 29.
RC TISSUE=Epididymis;
RX MEDLINE=99201788; PubMed=10101572;
RA Fan H.-Y., Miao S.-Y., Wang L.-F., Koide S.S.;
RT "Expression and characterization of an epididymis-specific gene."
RL Arch. Androl. 42:63-69(1999).
RN [2]
RP SEQUENCE OF 29-123 FROM N.A., AND SEQUENCE OF 29-58.
RC TISSUE=Epididymis;
RX MEDLINE=97040973; PubMed=8886263;
RA Xu W.D., Wang L.-F., Miao S.-Y., Zhao M., Fan H.-Y., Zong S.D.,
RA Wu Y.W., Shi X.Q., Koide S.S.;
RT "Identification of a rabbit epididymal protein gene."
RL Arch. Androl. 37:135-141(1996).
CC -!- FUNCTION: POSSIBLE FUNCTION IN SPERM MATURATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EPIDIDYMIS.

CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U26725; AAA66525.2; -.
DR HSP; Q9N0L8; 1TWP.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 2.
DR PRINTS; PR00003; 4DISULPHCORE.
DR ProDom; PD001224; WAP; 1.
DR SMART; SM00217; WAP; 2.
DR PROSITE; PS00317; 4-DISULFIDE_CORE; 2.
KW Repeat; Signal; Glycoprotein.
FT SIGNAL 1 28
FT CHAIN 29 123 MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4.
FT DOMAIN 33 75 WAP 1.
FT DOMAIN 77 123 WAP 2.
FT DISULFID 37 63 BY SIMILARITY.
FT DISULFID 46 67 BY SIMILARITY.
FT DISULFID 50 62 BY SIMILARITY.
FT DISULFID 56 71 BY SIMILARITY.
FT DISULFID 81 109 BY SIMILARITY.
FT DISULFID 92 113 BY SIMILARITY.
FT DISULFID 96 108 BY SIMILARITY.
FT DISULFID 102 118 BY SIMILARITY.
FT CARBOHYD 45 45 N-LINKED (GLCNAC...), (POTENTIAL).
SQ SEQUENCE 123 AA; 12803 MW; 99F9A649CCBB5B32 CRC64;

Query Match 30.7%; Score 100.5; DB 1; Length 123;
Best Local Similarity 39.6%; Pred. No. 0.00018;
Matches 21; Conservative 3; Mismatches 26; Indels 3; Gaps 1;

QY 4 PVKGPVSTKPGSCPIILIRCAMLPNPNRCLKDTDCPIKCKCGSCGMACFVP 56
DB 25 PVITGTGADKPGVCPQL---SADLNCTQDCRADQDCAENLKCRCRAGCSAICSIP 74

RESULT 11
WAP_MACEU
ID WAP_MACEU STANDARD; PRT; 191 AA.
AC Q9N0L8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Whey acidic protein precursor (tWAP).
GN WAP.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING.
RC TISSUE=Lactating mammary gland;
RX MEDLINE=20390063; PubMed=10801834;
RA Simpson K.J., Ranganathan S., Fisher J.A., Janssens P.A., Shaw D.C.,
RA Nicholas K.R.;
RT "The gene for a novel member of the whey acidic protein family encodes
RT three four-disulphide core domains and is asynchronously expressed
RT during lactation."
RL J. Biol. Chem. 275:23074-23081(2000).
CC -!- FUNCTION: COULD BE A PROTEASE INHIBITOR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: MILK-SPCIFIC; MAJOR PROTEIN COMPONENT OF MILK
CC WHEY.
CC -!- SIMILARITY: CONTAINS 3 WAP-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AJ005356; CAB90357.1; -;
 DR PIR; B47222; B47222.
 DR InterPro; IPR002221; WAP.
 DR Pfam; PF00095; wap; 3.
 DR PRINTS; PR00003; 4DISULPHCORE.
 DR ProDom; PD001224; WAP; 1.
 DR SMART; SM00217; WAP; 3.
 DR PROSITE; PS00317; 4 DISULFIDE CORE; 3.
 KW Milk; Whey; Protease inhibitor; Repeat; Signal; 3D-structure.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 191 WHEY ACIDIC PROTEIN.
 FT DOMAIN 25 71 WAP 1.
 FT DOMAIN 74 126 WAP 2.
 FT DOMAIN 129 175 WAP 3.
 FT DISULFID 29 59
 FT DISULFID 42 63
 FT DISULFID 46 58
 FT DISULFID 52 67
 FT DISULFID 78 114
 FT DISULFID 97 118
 FT DISULFID 101 113
 FT DISULFID 107 122
 FT DISULFID 133 163
 FT DISULFID 140 167
 FT DISULFID 150 162
 FT DISULFID 156 171
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 191 AA; 21133 MW; BF90B81DADB8E50D CRC64;

Query Match 30.7%; Score 100.5; DB 1; Length 191;
 Best Local Similarity 37.0%; Pred. No. 0.0026;
 Matches 20; Conservative 4; Mismatches 21; Indels 9; Gaps 2;

QY 8 PVSTKPGSCPILIRCAMLNPNNR-----CLKDTCPGIKKCCGSCGMACFVP 56
 DB 125 PVKAGRCRAVGTIC-----PEKKSFWHTCQRDDQCKENKCCSSACGRRCNP 174

RESULT 12
 ID KALM_CHICK STANDARD; PRT; 676 AA.
 AC P33005;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Anosmin 1 precursor (Kallmann syndrome protein homolog).
 GN KAL.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94010957; PubMed=8406507;
 RA Legouis R., Cohen-Salmon M., del Castillo I., Levilliers J.,
 RA Capy L., Morron J.-P., Petit C.;
 RT "Characterization of the chicken and quail homologues of the human
 RL Genomics 17:516-518(1993)
 CC -!- FUNCTION: May be an adhesion-like molecule with anti-protease
 CC activity.
 CC -!- SUBCELLULAR LOCATION: Secreted. Localized at cell surface (By
 CC similarity).
 CC -!- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L12144; AAA51435.1; -;
 DR PIR; B47222; B47222.
 DR HSP; P19957; 2REL.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR002221; WAP.
 DR Pfam; PF00041; fn3; 3.
 DR Pfam; PF00095; wap; 1.
 DR PRINTS; PR00003; 4DISULPHCORE.
 DR SMART; SM00060; FN3; 3.
 DR SMART; SM00217; WAP; 1.
 DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
 KW Cell adhesion; Glycoprotein; Serine protease inhibitor; Repeat;
 Signal.

FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 676 ANOSMIN 1.
 FT DOMAIN 22 115 "CYSTEINE BOX".
 FT DOMAIN 125 171 WAP.
 FT DOMAIN 176 280 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 281 397 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 398 536 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 537 657 FIBRONECTIN TYPE-III 4.
 FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 549 549 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 676 AA; 76375 MW; 3FAC7E082EA7E352 CRC64;

Query Match 30.7%; Score 100.5; DB 1; Length 676;
 Best Local Similarity 34.4%; Pred. No. 0.00074;
 Matches 21; Conservative 6; Mismatches 23; Indels 11; Gaps 2;

QY 3 EPVKGVPSTKPGSCPI-----ILIRCAMLNPNNRCLKDTCPGIKKCCGSCGMACFVP 56
 DB 116 EFLKYLVSVKQDCPAPEKASGFAACV-----ESCEADSECSGVKKCCSNGCGHTCQVP 170

QY 57 Q 57
 DB 171 K 171

RESULT 13

WAP_RAT
 ID WAP_RAT STANDARD; PRT; 137 AA.
 AC P01174;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Whey acidic protein precursor (Whey phosphoprotein) (WAP).
 GN WAP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85062841; PubMed=6095207;
 RA Campbell S.M., Rosen J.M., Hennighausen L.G., Strech-Jurk U.,
 RA Sippl A.E.;
 RT "Comparison of the whey acidic protein genes of the rat and mouse."
 RL Nucleic Acids Res. 12:8685-8697(1984).
 RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=82274212; PubMed=6896749;
RA Hennighausen L.G., Sippl A.E.;
RT "Comparative sequence analysis of the mRNAs coding for mouse and rat
RL whey protein.";
RL Nucleic Acids Res. 10:3733-3744 (1982).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82275050; PubMed=6955785;
RA Dandekar A.M., Robinson E.A., Appella E., Qasba P.K.;
RT "Complete sequence analysis of cDNA clones encoding rat whey
RT phosphoprotein: homology to a protease inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3987-3991 (1982).
CC -1- FUNCTION: COULD BE A PROTEASE INHIBITOR.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK
CC WHEY.
CC -1- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X01153; CAA25600.2; -
CC EMBL; X01154; CAA25600.2; JOINED.
CC EMBL; X01155; CAA25600.2; JOINED.
CC EMBL; X01156; CAA25600.2; JOINED.
CC EMBL; J00802; AAA42347.1; -
CC EMBL; J00801; AAA42346.1; -
CC FIC; A01228; WYRT.
CC FIC; A23879; A23879.
CC HSP; 046655; 1CJH.
CC InterPro: IPR002221; WAP.
CC Pfam: PF00095; wap; 2.
CC PRINTS; PR00003; 4DISULPHCORE.
CC PRODOM; PD001224; WAP; 1.
CC SMART; SM00217; WAP; 1.
CC PROSITE; PS00317; 4-DISULFIDE CORE; 2.
CC Milk; Whey; Protease inhibitor; Repeat; Phosphorylation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 137 WHEY ACIDIC PROTEIN.
FT DOMAIN 30 73 WAP 1 (ATYPICAL).
FT DOMAIN 79 127 WAP 2.
FT DISULFID 34 61 BY SIMILARITY.
FT DISULFID 45 65 BY SIMILARITY.
FT DISULFID 48 60 BY SIMILARITY.
FT DISULFID 54 69 BY SIMILARITY.
FT DISULFID 83 115 BY SIMILARITY.
FT DISULFID 96 119 BY SIMILARITY.
FT DISULFID 102 114 BY SIMILARITY.
FT DISULFID 108 123 BY SIMILARITY.
FT MOD RES 38 38 PHOSPHORYLATION (POTENTIAL).
FT MOD RES 40 40 PHOSPHORYLATION (POTENTIAL).
FT MOD RES 50 50 PHOSPHORYLATION (POTENTIAL).
FT MOD RES 103 103 PHOSPHORYLATION (POTENTIAL).
FT CONFLICT 4 4 S -> F (IN REF. 3).
FT CONFLICT 35 35 S -> P (IN REF. 2 AND 3).
FT CONFLICT 39 39 F -> S (IN REF. 2 AND 3).
FT CONFLICT 47 47 N -> K (IN REF. 2).
FT CONFLICT 68 68 S -> P (IN REF. 2 AND 3).
FT CONFLICT 99 99 D -> G (IN REF. 2).
FT CONFLICT 116 116 K -> N (IN REF. 2 AND 3).
FT CONFLICT 127 127 E -> K (IN REF. 2).
FT CONFLICT 129 129 K -> E (IN REF. 3).
FT CONFLICT 129 129 K -> D (IN REF. 2).
FT CONFLICT 134 134 I -> V (IN REF. 3).
SQ SEQUENCE 137 AA; 14827 MW; 1C2E8ADA9FD97949 CRC64;
Query Match 30.0%; Score 98; DB 1; Length 137;

Best Local Similarity 37.9%; Pred. No. 0.00037; Indels 6; Gaps 2;
Matches 22; Conservative
OY 6 KGPVS---TKPGSC---PIILIRCAMLNPPNRCUKDTDCPGIKKCCGSCGCMACFPQ 57
DB 70 KTPVNIQKAGRCFNPFIQIAAGPCPKDPKPCSIDSDCSGTMKCKCKNGCIMSCHDPE 127
RESULT 14
WDMN MOUSE STANDARD; PRT; 74 AA.
AC Q62477;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WDMN1 protein precursor.
GN WDMN1 OR EXPI.
OS Mus musculus (Mouse), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=FVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "neu and ras initiate murine mammary tumors that share genetic
RT markers generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine; TISSUE=Mammary gland;
RA Kho Y.J., Lee D.Y., Choi Y.J., Baik M.G.;
RT "Cloning and characterization of involution specific genes in bovine
RT mammary gland.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COULD HAVE PROTEINASE INHIBITING CAPACITY.
CC -1- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X93037; CAA63605.1; -
CC EMBL; AF159701; AAF76524.1; -
CC HSSP; Q9NOL8; 1TWP.
CC MGD; MGI:107506; Expi.
CC InterPro: IPR002221; WAP.
CC Pfam: PF00095; wap; 1.
CC PRINTS; PR00003; 4DISULPHCORE.
CC SMART; SM00217; WAP; 1.
CC PROSITE; PS00317; 4-DISULFIDE CORE; 1.
CC Protease inhibitor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 74 WDMN1 PROTEIN.
FT DOMAIN 29 73 WAP.
SQ SEQUENCE 74 AA; 7787 MW; 42BDCAD5B330B3D3 CRC64;
Query Match 29.5%; Score 96.5; DB 1; Length 74;
Best Local Similarity 37.7%; Pred. No. 0.00033;
Matches 20; Conservative 1; Mismatches 15; Indels 17; Gaps 2;
OY 12 KPGSCPIILIRCAMLNPP-----NRCLKDTDCPGIKKCCGSCGCMACFPV 56
DB 29 KPGACP-----KPPPRSFGTCDERTGTGDCSGNWKCCSNCCHACKRP 72
RESULT 15

Best Local Similarity 41.4%; Pred. No. 0.00067;
Matches 24; Conservative 3; Mismatches 22; Indels 9; Gaps 3;

Qv 4 PVKGPVSTKPGSCPIILIRCAMLNP----PNRCLKOTDCPGIKKCGSCGMACFVP 56
Db 72 PLLVPV-PKAGRCFWV---PAPLAPELCLEKNECSRDDQCRGNKKCFSSCAMRCLDP 125

RESULT 16
WFD2 HUMAN STANDARD; PRT; 124 AA.
ID Q14508; O96KJ1;
AC 15-JUN-1998 (Rel. 36, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major epididymis-specific protein E4 precursor (HE4) (Epididymal
DE secretory protein E4) (WAP four-disulfide core domain protein 2).
GN WPD2C OR HE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epididymis;
RX MEDLINE=92153963; PubMed=1686187;
RA Kirchhoff C., Habben L., Ivell R., Krull N.;
RT "A major human epididymis-specific cDNA encodes a protein with
RT sequence homology to extracellular proteinase inhibitors";
RL Biol. Reprod. 45:350-357(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehtvaslao M.H., Leversna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- FUNCTION: POSSIBLE FUNCTION IN SPERM MATURATION.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE EPIDIDYMAL DUCT,
CC PREDOMINANTLY WITHIN THE DISTAL SECTIONS.
CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)
CC or send an email to license@isb-sib.ch)

```

CC EMBL; X63187; CAA44869.1; --
DR EMBL; A18924; CAA01433.1; --
DR EMBL; A031663; CAB37641.1; --
DR HSP; Q9N0L8; ITWP.
DR InterPro; IPR002221; WAP.
DR SMART; SM00217; WAP; 3.
DR PROSITE; PS00317; 4 DISULFIDE_CORE; 2.
KW Repeat; Signal; Glycoprotein.
FT SIGNAL 1 21
FT CHAIN 22 124
FT DOMAIN 32 73
FT DOMAIN 76 124
FT DISULFID 36 62
FT DISULFID 45 66
FT DISULFID 49 61
FT DISULFID 55 70
FT DISULFID 80 110
FT DISULFID 93 114
FT DISULFID 97 109
FT DISULFID 103 119
FT CARBOHYD 44 44
FT CONFLICT 71 72
FT CONFLICT 101 101
SQ SEQUENCE 124 AA; 12993 MW; 9536B00B385259AD CRC64;

Query Match 28.4%; Score 93; DB 1; Length 124;
Best Local Similarity 40.4%; Pred. No. 0.0012;
Matches 19; Conservative 6; Mismatches 20; Indels 2; Gaps 2;

QY 12 KPGSCPIILIRCAMLN-PPNRCLKDTDCPGIKKCCGCG-MACFVP 56
DB 76 KEGSCFQVNIIPQLGLRQCDQVDSQCFQGMKCCRGCGKVCVTP 122

RESULT 17
KALM_HUMAN STANDARD; PRT; 680 AA.
AC P23352;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anosmin 1 precursor (Kallmann syndrome protein) (Adhesion molecule-
DE like X-linked).
GN KAL1 OR KAL OR ADMX OR KALIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92005720; PubMed=1913827;
RX Legouis R., Hardelin J.-P., Levilliers J., Claverie J.-M., Compain S.,
RA Wunderle V., Millaesau P., le Paslier D., Cohen D., Caterina D.,
RA Bougueleret L., Deleamarre-Van de Waal H., Lutfalla G., Weissbach J.,
RA Petit C.;
RT "The candidate gene for the X-linked Kallmann syndrome encodes a
RT protein related to adhesion molecules."
RL Cell 67:423-435 (1991).
RN [2]
RP REVISIONS.
RX MEDLINE=93265164; PubMed=1303284;
RX del Castillo I., Cohen-Salmon M., Blanchard S., Lutfalla G.,
RA Petit C.;
RT "Structure of the X-linked Kallmann syndrome gene and its homologous
RT pseudogene on the X chromosome."
RL Nat. Genet. 2:305-310 (1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92018217; PubMed=1922361;
RX Franco B., Guioli S., Pragliola A., Inceri B., Bardoni B.,

```

```

RA TAillon-Miller P., Brown C.J., Willard H.F., Lawrence C.,
RA Persico N.G., Camerino G., Ballabio A.;
RT "A gene deleted in Kallmann's syndrome shares homology with neural
RT cell adhesion and axonal path-finding molecules."
RL Nature 353:529-536 (1991).
RN [4]
RP SEQUENCE OF 1-71 FROM N.A.
RX MEDLINE=96069588; PubMed=7590336;
RA Cohen-Salmon M., Tronche F., del Castillo I., Petit C.;
RT "Characterization of the promoter of the human KAL gene, responsible
RT for the X-chromosome-linked Kallmann syndrome."
RL Gene 164:235-242 (1995).
RN [5]
RP CHARACTERIZATION.
RX PubMed=8832397;
RA Soussi-Yanicostas N., Hardelin J.-P., del Mar Arroyo-Jimenez M.,
RA Ardouin O., Legouis R., Levilliers J., Traincard F., Betton J.-M.,
RA Cabanie L., Petit C.;
RT "Initial characterization of anosmin-1, a putative extracellular
RT matrix protein synthesized by definite neuronal cell populations in
RT the central nervous system."
RL J. Cell Sci. 109:1749-1757 (1996).
RN [6]
RP VARIANT KAL1 LYS-267, AND VARIANT VAL-534.
RX MEDLINE=93278384; PubMed=8504298;
RA Hardelin J.-P., Levilliers J., Blanchard S., Carel J.-C.,
RA Leutenegger M., Pinard-Bertelletto J.-P., Bouloux P., Petit C.;
RT "Heterogeneity in the mutations responsible for X chromosome-linked
RT Kallmann syndrome."
RL Hum. Mol. Genet. 2:373-377 (1993).
RN [7]
RP VARIANT KAL1 LYS-514, AND VARIANT VAL-534.
RX MEDLINE=98251593; PubMed=9589672;
RA Maya-Nunez G., Zenteno J.C., Ulloa-Aguirre A., Kofman-Altaro S.,
RA Mendez J.P.;
RT "A recurrent missense mutation in the KAL gene in patients with
RT X-linked Kallmann's syndrome."
RL J. Clin. Endocrinol. Metab. 83:1650-1653 (1998).
CC -I- FUNCTION: May be an adhesion-like molecule with anti-protease
CC activity.
CC -I- SUBCELLULAR LOCATION: Secreted. Localized at cell surface.
CC -I- PTM: N-glycosylated.
CC -I- DISEASE: Defects in KAL1 are a cause of Kallmann syndrome (KAL1 OR
CC KS), a genetic disorder that associates hypogonadotropic
CC hypogonadism and anosmia. In this disease, the normal embryonic
CC migration of GnRH-synthesizing neurons from the olfactory placodes
CC to the hypothalamic region as well as the axonal extension of
CC olfactory neurons towards the forebrain are impaired.
CC -I- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M97252; AAA59202.1; --
DR EMBL; S60085; AAB20108.1; ALT_SEQ.
DR EMBL; X60299; CAA42841.1; --
DR EMBL; X82034; CAA57554.1; --
DR PIR; A40351; A40351.
DR PIR; S17982; S17982.
DR Genew; HGNC:6211; KAL1.
DR MIM; 308700; --
DR MIM; 147950; --
DR InterPro; IPR003961; FN III.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00095; wap; 1.

```



```

DR ProDom: PD001224; WAP: 1.
DR SMART: SM00217; WAP: 2.
DR PROSITE: PS00317; 4 DISULFIDE_CORE; 2.
KW Repeat; Signal; Glycoprotein.
FT SIGNAL 1 21
FT CHAIN 22 124 MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4.
FT DOMAIN 32 74 WAP 1.
FT DOMAIN 76 124 WAP 2.
FT DISULFID 36 62 BY SIMILARITY.
FT DISULFID 45 66 BY SIMILARITY.
FT DISULFID 49 61 BY SIMILARITY.
FT DISULFID 55 70 BY SIMILARITY.
FT DISULFID 80 110 BY SIMILARITY.
FT DISULFID 93 114 BY SIMILARITY.
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 119 BY SIMILARITY.
FT CARBOHYD 44 44 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 124 AA; 12951 MW; 15AAF315BA13958C CRC64;

Query Match 26.9%; Score 88; DB 1; Length 124;
Best Local Similarity 38.3%; Pred. No. 0.0042;
Matches 18; Conservative 7; Mismatches 20; Indels 2; Gaps 2;

QY 12 KPGSCPILIRCAMLN-PPNRLKDTDCPGIKKCEGSCG-MACFVP 56
Db 76 KEGSCPQVNTDFPQLGDCQDQVDSHCPGLLKCCYNGCKGVSCVTP 122

RESULT 20
EPPI MOUSE
ID EPPI MOUSE STANDARD; PRT; 134 AA.
AC Q9DA01;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Eppin precursor (Epididymal protease inhibitor) (Serine protease
DE inhibitor-like with Kunitz and WAP domains 1).
GN SPINLWI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C; TTSSUE=Testis, and Epididymis;
RA Sivashanmugam P., Hall S.H., Hamil K.G., French F.S., O'Rand M.G.,
RA Richardson R.T.;
RT "Characterization of monkey and mouse Eppin, a protease inhibitor from
RT epididymis and testis."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."

```

```

RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF346413; AAK31335.1; -.
DR EMBL; AK006296; BAB24514.1; -.
DR HSSP; P31713; 1SHP.
DR MGD; MGI:1922776; 1700024E17Rik.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR Pfam; PF00095; wap; 1.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00131; KU; 1.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE_CORE; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 134 EPPIN.
FT DOMAIN 29 73 WAP.
FT DOMAIN 77 127 BPTI/KUNITZ INHIBITOR.
FT DISULFID 33 61 BY SIMILARITY.
FT DISULFID 40 65 BY SIMILARITY.
FT DISULFID 48 60 BY SIMILARITY.
FT DISULFID 54 69 BY SIMILARITY.
FT DISULFID 77 127 BY SIMILARITY.
FT DISULFID 86 110 BY SIMILARITY.
FT DISULFID 102 123 BY SIMILARITY.
SQ SEQUENCE 134 AA; 15470 MW; DFE6B63D4D4C427F CRC64;

Query Match 26.9%; Score 88; DB 1; Length 134;
Best Local Similarity 35.7%; Pred. No. 0.0045;
Matches 20; Conservative 5; Mismatches 27; Indels 4; Gaps 2;

QY 5 VKGPVSTK---PGSCPILIRCAMLNPPNRLKDTDCPGIKKCEGSCGMACFVPQ 57
Db 19 VQGFSLADLLFFRCPRFRECEH-QERDLCTRDRCPKKCKCVFNCCKCLNPQ 73

RESULT 21
EPPI HUMAN
ID EPPI HUMAN STANDARD; PRT; 133 AA.
AC Q95925; Q9HD30; Q9SD7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Eppin precursor (Epididymal protease inhibitor) (Serine protease
DE inhibitor-like with Kunitz and WAP domains 1).
GN SPINLWI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Epididymis, and testis;
RX MEDLINE=21297183; PubMed=11404006;
RA Richardson R.T., Sivashanmugam P., Hall S.H., Hamil K.G., Moore P.A.,
RA Ruben S.M., French F.S., O'Rand M.G.;
RT "Cloning and sequencing of human Eppin: a novel family of protease

```

RT inhibitors expressed in the epididymis and testis.;"

RL Gene 270:93-102(2001).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RA Stavrides G.S., Huckle E.J., Deloukas P.;

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosht K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA Levasialho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,

RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W., Williams S.A.,

RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.;"

RL Nature 414:865-871(2001).

CC -1- SUBCELLULAR LOCATION: Secreted (Potential).

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.

CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF286370; AAG00548.1; -

DR EMBL; AF286369; AAG00547.1; -

DR EMBL; AF286368; AAG00546.1; -

DR EMBL; AL118493; CAB56343.1; -

DR EMBL; AL031663; CAB37635.1; -

DR EMBL; AL031663; CAB36265.1; -

DR HSSP; P00974; IBPI.

DR Genew; HGNC:15932; SPINLW1.

DR InterPro; IPR002223; Kunitz_BPTI.

DR InterPro; IPR002221; WAP.

DR Pfam; PF00014; Kunitz_BPTI; 2.

DR Pfam; PF00095; wap; 2.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR SMART; SM00217; WAP; 1.

DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.

DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.

DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.

KW Serine protease inhibitor; Signal; Alternative splicing.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 133 EPPIN.

FT DOMAIN 29 73 WAP.

FT DISULFID 77 127 BPTI/KUNITZ INHIBITOR.

FT DISULFID 33 61 BY SIMILARITY.

FT DISULFID 40 65 BY SIMILARITY.

FT DISULFID 48 60 BY SIMILARITY.

FT DISULFID 54 69 BY SIMILARITY.

FT DISULFID 77 127 BY SIMILARITY.

FT DISULFID 86 110 BY SIMILARITY.

FT DISULFID 102 123 BY SIMILARITY.

FT VARSPLIC 1 31 MGSSGLLSLVLFVILLANVQPGGLTDWLFPR -> MLSKAH

FT CCKTALSUG (IN ISOFORM 2).

SQ SEQUENCE 133 AA; 15284 MW; F7831B203366D9DC CRC64;

Query Match 26.0%; Score 85; DB 1; Length 133;

Best Local Similarity 40.4%; Pred. No. 0.0094;

Matches 21; Conservative 2; Mismatches 25; Indels 4; Gaps 2;

QY 5 VQGPVSTK---PGSCPILIRCAMLPNRCCLKDTDCFGIKKCCGSCGMAC 53

Db 19 VQGPGLTDWLFPRPCPKIRECE-FOERDVCTKDRQCQDNKKCCVFCGKKC 69

RESULT 22

EPPI_MACMU

ID EPPI_MACMU STANDARD; PRT; 133 AA.

AC Q9BDL1;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Eppin precursor (Epididymal protease inhibitor) (serine protease

DE inhibitor-like with Kunitz and WAP domains 1).

GN SPINLW1.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

OC Cercopitheinae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis, and Epididymis;

RA Sivashanmugam P., Hall S.H., Hamil K.G., French F.S., O'Rand M.G.,

RA Richardson R.F.;

RT "Characterization of monkey and mouse Eppin, a protease inhibitor from

RT epididymis and testis."

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Secreted (Potential).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.

CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF346414; AAK31336.1; -

DR HSSP; P00974; IBPI.

DR InterPro; IPR002223; Kunitz_BPTI.

DR InterPro; IPR002221; WAP.

DR Pfam; PF00014; Kunitz_BPTI; 1.

DR Pfam; PF00095; wap; 1.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR SMART; SM00217; WAP; 1.

DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.

DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.

KW Serine protease inhibitor; Signal.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 133 EPPIN.

FT DOMAIN 29 73 WAP.
 FT DOMAIN 77 127 BPT1/KUNITZ INHIBITOR.
 FT DISULFID 33 61 BY SIMILARITY.
 FT DISULFID 40 65 BY SIMILARITY.
 FT DISULFID 48 60 BY SIMILARITY.
 FT DISULFID 54 69 BY SIMILARITY.
 FT DISULFID 77 127 BY SIMILARITY.
 FT DISULFID 86 110 BY SIMILARITY.
 FT DISULFID 102 123 BY SIMILARITY.
 SQ SEQUENCE 133 AA; 15279 MW; 433A894639A35E9 CRC64;
 Query Match 26.0%; Score 85; DB 1; Length 133;
 Best Local Similarity 38.5%; Pred. No. 0.0094;
 Matches 20; Conservative 3; Mismatches 25; Indels 4; Gaps 2;
 QY 5 VKGPVSTK---PGSCPIILIRCAMLPNPNRCLKDTDCPGIKKCCGSCGMAC 53
 DB 19 VQPGGLTDWLFRRCPRTIRECE-FRERDVCTHRCQCPDNKKCCVFCGKKC 69
 RESULT 23
 WFD1 RAT
 ID WFD1 RAT STANDARD; PRT; 212 AA.
 AC O70280;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WAP four-disulfide core domain protein 1 protein precursor (Prostate
 stromal protein ps20) (ps20 growth inhibitor).
 GN WFDCL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Prostate;
 RX MEDLINE=98136158; PubMed=9468514;
 RA Larsen M., Ressler S.J., Lu B., Gerdes M.J., McBride L., Dang T.D.,
 RA Rowley D.R.;
 RT "Molecular cloning and expression of ps20 growth inhibitor. A novel
 WAP-type 'four-disulfide core' domain protein expressed in smooth
 muscle.";
 RL J. Biol. Chem. 273:4574-4584 (1998).
 RN [2]
 RP PARTIAL SEQUENCE OF 27-54, AND CHARACTERIZATION.
 RX PubMed=7665628;
 RA Rowley D.R., Dang T.D., Larsen M., Gerdes M.J., McBride L., Lu B.;
 RT "Purification of a novel protein (ps20) from urogenital sinus
 mesenchymal cells with growth inhibitory properties in vitro.";
 RL J. Biol. Chem. 270:22058-22065 (1995).
 CC -!- FUNCTION: Has growth inhibitory activity.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Vascular smooth muscle and prostate.
 CC Periacinar ring.
 CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; AF037272; AAC40055.1; -
 DR InterPro; IPR002221; WAP.
 DR Pfam; PF00095; wap; 1.
 DR SMART; SM00217; WAP; 1.
 DR PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
 KW SIGNAL.
 FT SIGNAL 1 26 WAP FOUR-DISULFIDE CORE DOMAIN PROTEIN 1
 FT CHAIN 27 212

FT DOMAIN 54 99 WAP.
 FT DISULFID 58 88 BY SIMILARITY.
 FT DISULFID 70 92 BY SIMILARITY.
 FT DISULFID 75 87 BY SIMILARITY.
 FT DISULFID 81 96 BY SIMILARITY.
 SQ SEQUENCE 212 AA; 23230 MW; 2BB054DE9D8460A5 CRC64;
 Query Match 24.9%; Score 81.5; DB 1; Length 212;
 Best Local Similarity 30.4%; Pred. No. 0.033;
 Matches 14; Conservative 6; Mismatches 15; Indels 11; Gaps 1;
 QY 8 PVSTKPGSCPIILIRCAMLPNPNRCLKDTDCPGIKKCCGSCGMAC 53
 DB 62 PRTLPGAC-----QATRCQSDSECPHRRCCYNGCAYAC 96
 RESULT 24
 WFD1 MOUSE
 ID WFD1 MOUSE STANDARD; PRT; 211 AA.
 AC Q9ESH5;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WAP four-disulfide core domain protein 1 protein precursor (Prostate
 stromal protein ps20) (ps20 growth inhibitor).
 GN WFDCL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=20424791; PubMed=10967136;
 RA Larsen M., Ressler S.J., Gerdes M.J., Lu B., Byron M., Lawrence J.B.,
 RA Rowley D.R.;
 RT "The WFDCL gene encoding ps20 localizes to 16q24, a region of LOH in
 multiple cancers.";
 RL Mamm. Genome 11:767-773 (2000).
 RN [2]
 RP SEQUENCE OF 4-211 FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Has growth inhibitory activity (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; AF170446; AAC21086.1; -
 DR EMBL; AF169632; AAC21086.1; JOINED.
 DR EMBL; BC025850; AAH25850.1; -
 DR InterPro; IPR002221; WAP.
 DR Pfam; PF00095; wap; 1.
 DR SMART; SM00217; WAP; 1.
 DR PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
 KW SIGNAL.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 WAP FOUR-DISULFIDE CORE DOMAIN PROTEIN 1
 FT DOMAIN 53 98 WAP.
 FT DISULFID 57 87 BY SIMILARITY.
 FT DISULFID 69 91 BY SIMILARITY.
 FT DISULFID 74 86 BY SIMILARITY.
 FT DISULFID 80 95 BY SIMILARITY.

FT CONFLICT 132 132 MISSING (IN REF. 2).
 SQ SEQUENCE 211 AA; 23079 MW; 605E7E6B0FC64F55 CRC64;
 Query Match 24.6%; Score 80.5; DB 1; Length 211;
 Best Local Similarity 30.4%; Pred. No. 0.042;
 Matches 14; Conservative 6; Mismatches 15; Indels 11; Gaps 1;
 QY 8 PVSITKPGSCPILLRCAMLPNPKLKDTPGIGKKCEGSCGMAC 53
 Db 61 PRTLPFGAC-----QATRCQADSECPHRRCCYNGCAVAC 95
 RESULT 25
 WAP_MOUSE STANDARD; PRT; 134 AA.
 AC P01173; P70230; Q61023;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE WAP: whey acidic protein precursor (WAP).
 GN WAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85062841; PubMed=6095207;
 RA Campbell S.M., Rosen J.M., Hennighausen L.G., Strech-Jurk U.,
 RA Sippl A.E.;
 RT "Comparison of the whey acidic protein genes of the rat and mouse";
 RL Nucleic Acids Res. 12:8685-8697(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82196900; PubMed=6896234;
 RA Hennighausen L.G., Sippl A.E.;
 RT "Whey acidic protein is a novel member of the family of 'four-
 RT disulfide core' proteins";
 RL Nucleic Acids Res. 10:2677-2684 (1982).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GR;
 RA Hennighausen L.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP CHARACTERIZATION.
 RC STRAIN=YBR;
 RX MEDLINE=82052974; PubMed=6975276;
 RA Piletz J.E., Heinlen M., Ganschow R.E.;
 RT "Biochemical characterization of a novel whey protein from murine
 RT milk";
 RL J. Biol. Chem. 256:11509-11516(1981).
 CC -1- FUNCTION: COULD BE A PROTEASE INHIBITOR. MAY PLAY AN IMPORTANT
 CC ROLE IN MAMMARY GLAND DEVELOPMENT AND TISSUE REMODELING.
 CC -1- SUBCELLULAR LOCATION: Secreted
 CC -1- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK
 CC WHEY.
 CC -1- PTM: NO PHOSPHATE OR CARBOHYDRATE BINDING COULD BE DETECTED;
 CC HOWEVER, BOTH CHOLESTEROL AND TRIGLYCERIDE ARE ASSOCIATED WITH THE
 CC MOUSE PROTEIN.
 CC -1- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X01157; CAA25604.1; -;
 DR EMBL; X01158; CAA25604.1; JOINED.
 DR EMBL; X01159; CAA25604.1; JOINED.

DR EMBL; X01160; CAA25604.1; JOINED.
 DR EMBL; U00856; AAA24224.1; -;
 DR EMBL; U38816; AAA91321.1; -;
 DR PIR; A01227; WYMS.
 DR PIR; B23879; B23879.
 DR HSSP; O46655; ICJH.
 DR MGD; MGI:98943; Wap.
 DR InterPro; IPR002221; WAP.
 DR Pfam; PF00095; WAP; 2.
 DR PRINTS; PR00003; 4DISULPHCORE.
 DR SMART; SM00217; WAP; 1.
 DR PROSITE; PS00317; 4-DISULFIDE_CORE; 2.
 KW Milk; Whey; Protease inhibitor; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 134 WHEY ACIDIC PROTEIN.
 FT DOMAIN 24 72 WAP 1 (ATYPICAL).
 FT DOMAIN 79 128 WAP 2.
 FT DISULFID 83 116 BY SIMILARITY.
 FT DISULFID 97 120 BY SIMILARITY.
 FT DISULFID 103 115 BY SIMILARITY.
 FT DISULFID 109 124 BY SIMILARITY.
 FT CONFLICT 2 2 R -> S (IN REF. 1).
 FT CONFLICT 11 11 L -> R (IN REF. 3).
 FT CONFLICT 35 35 Q -> P (IN REF. 2).
 FT CONFLICT 63 63 G -> V (IN REF. 1).
 FT CONFLICT 63 63 G -> R (IN REF. 2).
 FT CONFLICT 87 87 L -> S (IN REF. 1).
 FT CONFLICT 100 100 K -> Q (IN REF. 2).
 SQ SEQUENCE 134 AA; 14423 MW; C12B2544877ECA80 CRC64;
 Query Match 24.3%; Score 79.5; DB 1; Length 134;
 Best Local Similarity 34.5%; Pred. No. 0.037;
 Matches 20; Conservative 4; Mismatches 27; Indels 7; Gaps 2;
 QY 6 KQPV-----TKPGSCPILLRCAMLPN-----NRCLKDTPGIGKKCEGSCGMACFVP 56
 Db 70 KTPVNGVFKAGFCFVNLQTISSGTGCPMKIEGSSDRECSGNMKCCNVDCVMTCTPP 127
 RESULT 26
 WFD1_HUMAN
 ID WFD1_HUMAN STANDARD; PRT; 220 AA.
 AC Q9HC57; Q9HAU1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WAP four-disulfide core domain protein 1 protein precursor (Prostate
 DE stromal protein ps20) (ps20 growth inhibitor).
 GN WFDCL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=20424791; PubMed=10967136;
 RA Larsen M., Ressler S.J., Gerdes M.J., Lu B., Byron M., Lawrence J.B.,
 RA Rowley D.R.;
 RT "The WFDCL gene encoding ps20 localizes to 16q24, a region of LOH in
 RT multiple cancers";
 RL Mamm. Genome 11:767-773(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX Ong C.K., Ng C.Y., Lim K.B., Chan T.W.M.G., Huynh H.;
 RT "Molecular cloning and characterization of the human ps20 protein in
 RT human uterus";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon, Kidney, and Stomach;
 RA Strausberg R.;

DR Pfam: PF00014; Kunitz_BPTI; 1.
 DR Pfam: PF00095; wap; 1.
 DR PRINTS: PR00003; 4DISULPHCORE.
 DR PRINTS: PR00759; BASICPTASE.
 DR ProDom: PD000222; Kunitz_BPTI; 1.
 DR SMART: SM00131; KU; 1.
 DR SMART: SM00217; WAP; 1.
 DR PROSITE: PS00317; 4DISULFIDE CORE; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS0279; BPTI_KUNITZ_2; 1.
 KW Serine protease inhibitor.
 FT DOMAIN 8 58 BPTI/KUNITZ INHIBITOR.
 FT MOD_RES 63 105 WAP.
 FT DISULFID 1 1 PYROLIDONE CARBOXYLIC ACID.
 FT DISULFID 8 58 BY SIMILARITY.
 FT DISULFID 17 41 BY SIMILARITY.
 FT DISULFID 33 54 BY SIMILARITY.
 FT DISULFID 67 92 BY SIMILARITY.
 FT DISULFID 76 97 BY SIMILARITY.
 FT DISULFID 80 93 BY SIMILARITY.
 FT DISULFID 86 101 BY SIMILARITY.
 FT ACT_SITE 18 19 REACTIVE BOND (TRYPSIN).
 SQ SEQUENCE 110 AA; 11916 MW; 269436243813418E CRC64;
 Query Match 22.8%; Score 74.5; DB 1; Length 110;
 Best Local Similarity 34.8%; Pred. No. 0.11; 15; Indels 11; Gaps 2;
 Matches 16; Conservative 4; Mismatches 11; Gaps 2;
 QY 12 KPGSCPILIRCAMLNPNRCL-----KDTCPGPKKCGSGCGMAC 53
 DB 63 RGVCP-----KTSFGICLHCDSDSCKCKGQKCFDGGVYC 101
 RESULT 29
 ITRI MOMCO STANDARD; PRT; 34 AA.
 AC P82408:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Trypsin inhibitor I (MCOTI-I).
 OS Momordica cochinchinensis (Spiny bitter cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
 OX NCBI_TaxID=3674;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Seed;
 RX MEDLINE=20263490; PubMed=10801322;
 RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,
 RA Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;
 RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an
 RT atypical macrocyclic structure";
 RL Biochemistry 39:5722-5730(2000).
 CC -!- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT
 CC DEFENSE MECHANISM.
 CC -!- PTM: THIS IS A CYCLIC PEPTIDE.
 CC -!- MASS SPECTROMETRY: MW=3480.7; MW ERR=0.3; METHOD=Electrospray.
 CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
 CC INHIBITORS.
 DR HSSP: P12071; 2ETI.
 DR InterPro: IPR000737; Squash.
 DR Pfam: PF00299; squash; 1.
 DR PRINTS: PR00293; SQUASHINHTR.
 DR ProDom: PD003401; Squash; 1.
 DR SMART: SM00286; PTI; 1.
 DR PROSITE: PS00286; SQUASH_INHIBITOR; 1.
 KW Serine protease inhibitor.
 FT ACT_SITE 10 11 REACTIVE BOND.
 FT DISULFID 8 25 BY SIMILARITY.
 FT DISULFID 15 27 BY SIMILARITY.
 FT DISULFID 21 33 BY SIMILARITY.

SQ SEQUENCE 34 AA; 3505 MW; 8F7DOB4C162C935A CRC64;
 Query Match 20.5%; Score 67; DB 1; Length 34;
 Best Local Similarity 42.1%; Pred. No. 0.28;
 Matches 16; Conservative 2; Mismatches 10; Indels 10; Gaps 2;
 QY 14 GSCPIILIRCAMLNPNRCLKDTCPGKIKK-CEGSCG 50
 DB 6 GVCPKIL-----QRCRSDCFGACICRNGYCG 34
 RESULT 30
 FBN2 HUMAN STANDARD; PRT; 2911 AA.
 ID AC P35556:
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrillin 2 precursor.
 GN FBN2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94165150; PubMed=8120105;
 RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguinetti C.,
 RA Bonadio J., Mecham R.P., Ramirez F.;
 RT "Structure and expression of fibrillin-2, a novel microfibrillar
 RT component preferentially located in elastic matrices";
 RL J. Cell Biol. 124:855-863(1994).
 RN [2]
 RP SEQUENCE OF 752-1505 FROM N.A.
 RX MEDLINE=91304567; PubMed=4852206;
 RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
 RA Tsipouras P., Ramirez F., Hollister D.;
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to
 RT two different fibrillin genes";
 RL Nature 352:330-334(1991).
 RN [3]
 RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
 RX MEDLINE=96083599; PubMed=7493032;
 RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
 RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
 RT congenital contractual arachnodactyly";
 RL Nat. Genet. 11:456-458(1995).
 RN [4]
 RP VARIANTS CCA HIS-1114.
 RX MEDLINE=98407789; PubMed=9737771;
 RA Babcock D., Ganser C., Francke U., Maslen C.;
 RT "A single mutation that results in an asp-co-his substitution and
 RT partial exon skipping in a family with congenital contractual
 RT arachnodactyly";
 RL Hum. Genet. 103:22-28(1998).
 RN [5]
 RP VARIANTS CCA PHE-1141 AND TRP-1252.
 RX MEDLINE=20259236; PubMed=10797416;
 RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
 RA Godfrey M.;
 RT "Two novel fibrillin-2 mutations in congenital contractual
 RT arachnodactyly";
 RL Am. J. Med. Genet. 92:7-12(2000).
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
 CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
 CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
 CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
 CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
 CC AORTA AND THE EYES.
 CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.


```

Db 2018 ECVLPVSCSGTQNLGSPRC--ICPPGVKSENCIDINECDPNTCLFGSCTNTP 2075
Qy 50 -GMACFPV 56
Db 2076 GGFQCLCP 2083

RESULT 31
ID BAR3_CHITE STANDARD; PRT; 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Balbani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidea; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1699777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbani ring 3 gene in Chironomus tentans has a diverged
RT repetitive structure split by many introns.";
RL J. Mol. Biol. 211:331-349(1990).
CC -|- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -|- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52263; CAA36506.1; -
DR PIR; S08167; S08167.
DR HSSP; P15358; LSKZ.
DR InterPro; IPR004153; CXXC_repeat.
DR Pfam; PF03128; CXXC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 3420282521B0815 CRC64;

Query Match 19.9%; Score 65; DB 1; Length 1700;
Best Local Similarity 32.7%; Pred. No. 11;
Matches 17; Conservative 7; Mismatches 14; Indels 14; Gaps 4;

Qy 13 PGSCPIILTRCAMLNPPN--RLCKDT---DCPGIKKC-----CGSGCGMAC 53
Db 267 PNTCE---CGCAQLNCPDNKANKETCQCECKEVKCKNGQGVFCKDSCSCVC 315

RESULT 32
ID ITR3_MOMCO STANDARD; PRT; 30 AA.
AC P82410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Trypsin inhibitor III (MCOTI-III).
DE Momordica cochinchinensis (Spiny bitter cucumber).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
DE eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3674;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Seed;
RX MEDLINE=20263490; PubMed=10801322;
RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,
RA Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;
RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an
RT atypical macrocyclic structure.";
RL Biochemistry 39:5722-5730(2000).
CC -|- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT
CC DEFENSE MECHANISM.
CC -|- MASS SPECTROMETRY: MW=3379.6; MW ERR=0.5; METHOD=Electrospray.
CC -|- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR HSSP; P01074; 1BXJ.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHBTR.
DR PRODOM; PD003401; Squash; 1.
DR SMART; SM00286; PTI; 1.
DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
KW Serine protease inhibitor.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT ACT_SITE 6 7 REACTIVE BOND.
FT DISULFID 4 21 BY SIMILARITY.
FT DISULFID 11 23 BY SIMILARITY.
FT DISULFID 17 29 BY SIMILARITY.
SQ SEQUENCE 30 AA; 3403 MW; D5E13476B1A9AA6E CRC64;

Query Match 19.6%; Score 64; DB 1; Length 30;
Best Local Similarity 40.5%; Pred. No. 0.53;
Matches 15; Conservative 4; Mismatches 8; Indels 10; Gaps 2;

Qy 15 SCPILIRCAMLNPRLCKDTDCPGIKKCCE-CSCG 50
Db 3 ACPRL-----KKCRDSDCPGECICKENGYCG 30

RESULT 33
ID ITR2_MOMCO STANDARD; PRT; 34 AA.
AC P82409;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trypsin inhibitor II (MCOTI-II).
DE Momordica cochinchinensis (Spiny bitter cucumber).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
DE eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3674;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Seed;
RX MEDLINE=20263490; PubMed=10801322;
RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,
RA Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;
RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an
RT atypical macrocyclic structure.";
RL Biochemistry 39:5722-5730(2000).
CC -|- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT
CC DEFENSE MECHANISM.
CC -|- PTM: A CYCLIC SUCCINIMIDE MAY BE FORMED BETWEEN ASP-4 AND GLY-5
CC RESIDUES WHICH CAN BE TRANSFORMED TO A BETA-ASPARTYL BOND. SO
CC THERE EXISTS THREE ISOFORMS OF MCOTI-II, NORMAL, WITH SUCCINIMIDE
CC OR WITH A BETA-ASPARTYL BOND.
CC -|- PTM: THIS IS A CYCLIC PEPTIDE.

```

CC -1- MASS SPECTROMETRY: MW=3453; MW_ERR=0.2; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR HSSP: P01074; 1BXJ.
DR InterPro: IPR000737; Squash.
DR Pfam: PF00299; squash.1.
DR PRINTS: PR00293; SQUASHINHBTR.
DR PRODOM: PD003401; Squash; 1.
DR SMART: SM00286; PTI; 1.
DR PROSITE: PS00286; SQUASH INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT_SITE 10 11 REACTIVE_BOND.
FT DISULFID 8 25 BY SIMILARITY.
FT DISULFID 15 27 BY SIMILARITY.
FT DISULFID 21 33 BY SIMILARITY.
SQ SEQUENCE 34 AA; 3477 MW; 8F7D0B4C048BB93A CRC64;
Query Match 19.68; Score 64; DB 1; Length 34;
Best Local Similarity 39.5%; Pred. No. 0.58; Indels 10; Gaps 2;
Matches 15; Conservative 3; Mismatches 10; Indels 10; Gaps 2;
Qy 14 GSCPIILIRCAMLPNPPNRLKDTDCPGIKKC-CRSGSCG 50
Db 6 GVCPEKIL-----KKCRDSDCPGACICRGNGYCG 34
RESULT 34
FBN2 MOUSE STANDARD; PRT; 2907 AA.
ID FBN2 MOUSE STANDARD; PRT; 2907 AA.
AC Q61555; Q63957;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2 OR FBN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95263670; PubMed=7744963;
RA Zhang H., Hu W., Ramirez F.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
RT of extracellular microfibrils.";
RL J. Cell Biol. 129:1165-1176(1995).
RN [2]
RP SEQUENCE OF 210-317 FROM N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
RT mouse chromosomes 2 and 18.";
RL Genomics 18:667-672(1993).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; L39790; AAA74908.1; -;
CC ENBL; S69359; AAC60685.1; -;
CC HSSP; P35555; 1EMN.
CC MGD; MGI:95490; Fbn2.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF Ca.
DR InterPro: IPR001438; EGF II.
DR InterPro: IPR002212; Fibril-assoc.
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS: PR00010; EGFBL00D.
DR SMART: SM00179; EGF CA; 43.
DR SMART: SM00001; EGF-like; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 43.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS00022; EGF_2; 36.
DR PROSITE: PS01186; EGF CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
Repeat; Signal; Multigene family.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2907 FIBRILLIN 2.
FT DOMAIN 111 142 EGF-LIKE 1.
FT DOMAIN 145 176 EGF-LIKE 2.
FT DOMAIN 176 208 EGF-LIKE 3.
FT DOMAIN 276 317 EGF-LIKE 4.
FT DOMAIN 318 359 EGF-LIKE 5.
FT REPEAT 360 426 TGFBP 1.
FT DOMAIN 487 527 EGF-LIKE 6.
FT DOMAIN 528 567 EGF-LIKE 7.
FT DOMAIN 568 609 EGF-LIKE 8.
FT DOMAIN 610 650 EGF-LIKE 9.
FT DOMAIN 651 691 EGF-LIKE 10.
FT REPEAT 692 760 TGFBP 2.
FT DOMAIN 761 802 EGF-LIKE 11.
FT DOMAIN 803 844 EGF-LIKE 12.
FT DOMAIN 845 883 EGF-LIKE 13.
FT DOMAIN 948 989 EGF-LIKE 14.
FT REPEAT 990 1065 TGFBP 3.
FT DOMAIN 1066 1107 EGF-LIKE 15.
FT DOMAIN 1108 1150 EGF-LIKE 16.
FT DOMAIN 1151 1192 EGF-LIKE 17.
FT DOMAIN 1193 1234 EGF-LIKE 18.
FT DOMAIN 1235 1275 EGF-LIKE 19.
FT DOMAIN 1276 1317 EGF-LIKE 20.
FT DOMAIN 1318 1359 EGF-LIKE 21.
FT DOMAIN 1360 1400 EGF-LIKE 22.
FT DOMAIN 1401 1441 EGF-LIKE 23.
FT DOMAIN 1442 1483 EGF-LIKE 24.
FT DOMAIN 1484 1524 EGF-LIKE 25.
FT DOMAIN 1525 1565 EGF-LIKE 26.
FT REPEAT 1566 1642 TGFBP 4.
FT DOMAIN 1643 1684 EGF-LIKE 27.
FT DOMAIN 1685 1726 EGF-LIKE 28.
FT REPEAT 1727 1800 TGFBP 5.
FT DOMAIN 1801 1842 EGF-LIKE 29.
FT DOMAIN 1843 1884 EGF-LIKE 30.
FT DOMAIN 1885 1926 EGF-LIKE 31.
FT DOMAIN 1927 1965 EGF-LIKE 32.
FT DOMAIN 1966 2008 EGF-LIKE 33.
FT DOMAIN 2009 2048 EGF-LIKE 34.
FT DOMAIN 2049 2090 EGF-LIKE 35.
FT REPEAT 2091 2163 TGFBP 6.
FT DOMAIN 2164 2205 EGF-LIKE 36.
FT DOMAIN 2206 2245 EGF-LIKE 37.
FT DOMAIN 2246 2286 EGF-LIKE 38.
FT DOMAIN 2287 2330 EGF-LIKE 39.
FT DOMAIN 2331 2372 EGF-LIKE 40.
FT REPEAT 2373 2441 TGFBP 7.
FT DOMAIN 2442 2483 EGF-LIKE 41.
FT DOMAIN 2484 2524 EGF-LIKE 42.
FT DOMAIN 2525 2566 EGF-LIKE 43.
FT DOMAIN 2567 2607 EGF-LIKE 44.
FT DOMAIN 2607 2646 EGF-LIKE 45.
FT DOMAIN 2647 2687 EGF-LIKE 46.
FT DOMAIN 2688 2727 EGF-LIKE 47.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 125 130 BY SIMILARITY.

```
FT DISULFID 132 141 BY SIMILARITY.
FT DISULFID 149 159 BY SIMILARITY.
FT DISULFID 153 164 BY SIMILARITY.
FT DISULFID 166 170 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 184 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.
FT DISULFID 280 292 BY SIMILARITY.
FT DISULFID 287 301 BY SIMILARITY.
FT DISULFID 303 316 BY SIMILARITY.
FT DISULFID 322 334 BY SIMILARITY.
FT DISULFID 329 343 BY SIMILARITY.
FT DISULFID 345 358 BY SIMILARITY.
FT DISULFID 491 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 526 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 537 551 BY SIMILARITY.
FT DISULFID 553 566 BY SIMILARITY.
FT DISULFID 572 584 BY SIMILARITY.
FT DISULFID 579 593 BY SIMILARITY.
FT DISULFID 595 608 BY SIMILARITY.
FT DISULFID 614 625 BY SIMILARITY.
FT DISULFID 620 634 BY SIMILARITY.
FT DISULFID 636 649 BY SIMILARITY.
FT DISULFID 655 666 BY SIMILARITY.
FT DISULFID 661 675 BY SIMILARITY.
FT DISULFID 677 690 BY SIMILARITY.
FT DISULFID 765 777 BY SIMILARITY.
FT DISULFID 772 786 BY SIMILARITY.
FT DISULFID 788 801 BY SIMILARITY.
FT DISULFID 807 819 BY SIMILARITY.
FT DISULFID 814 828 BY SIMILARITY.
FT DISULFID 830 843 BY SIMILARITY.
FT DISULFID 849 859 BY SIMILARITY.
FT DISULFID 854 868 BY SIMILARITY.
FT DISULFID 870 883 BY SIMILARITY.
FT DISULFID 952 964 BY SIMILARITY.
FT DISULFID 959 973 BY SIMILARITY.
FT DISULFID 975 988 BY SIMILARITY.
FT DISULFID 1070 1082 BY SIMILARITY.
FT DISULFID 1077 1091 BY SIMILARITY.
FT DISULFID 1093 1106 BY SIMILARITY.
FT DISULFID 1112 1124 BY SIMILARITY.
FT DISULFID 1119 1133 BY SIMILARITY.
FT DISULFID 1135 1149 BY SIMILARITY.
FT DISULFID 1155 1167 BY SIMILARITY.
FT DISULFID 1162 1176 BY SIMILARITY.
FT DISULFID 1178 1191 BY SIMILARITY.
FT DISULFID 1197 1209 BY SIMILARITY.
FT DISULFID 1204 1218 BY SIMILARITY.
FT DISULFID 1220 1233 BY SIMILARITY.
FT DISULFID 1239 1250 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1261 1274 BY SIMILARITY.
FT DISULFID 1280 1292 BY SIMILARITY.
FT DISULFID 1287 1301 BY SIMILARITY.
FT DISULFID 1303 1316 BY SIMILARITY.
FT DISULFID 1322 1334 BY SIMILARITY.
FT DISULFID 1329 1343 BY SIMILARITY.
FT DISULFID 1345 1358 BY SIMILARITY.
FT DISULFID 1364 1377 BY SIMILARITY.
FT DISULFID 1371 1386 BY SIMILARITY.
FT DISULFID 1388 1399 BY SIMILARITY.
FT DISULFID 1405 1418 BY SIMILARITY.
FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1469 1482 BY SIMILARITY.
FT DISULFID 1488 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.

Query Match 19.3%; Score 63; DB 1; Length 2907;
Best Local Similarity 30.9%; Pred. No. 29;
Matches 21; Conservative 7; Mismatches 24; Indels 16; Gaps 4;

QY 3 BPVKGPSVTKPGSCPII--LIRCAMLNPPNRCCLKDTCGPKKCC-----GSC---- 49
Db 2012 ECVALPGSCSPGTQCNLEGSFRC--ICPPGYEVSENCIDNECDENICLFGSCTWTP 2069
QY 50 -QMAGFVP 56
Db 2070 GGFQICP 2077

RESULT 35
ADEN_ADEB7
ID ADEN_ADEB7 STANDARD; PRT; 202 AA.
AC P19151;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenain (EC 3.4.22.39) (Endoprotease) (Late L3 23 kDa protein).
OS Bovine adenovirus type 7 (Mastadenovirus bos7).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Atadenovirus.
OX NCBI_TaxID=10511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91016864; PubMed=2216744;
RA Cai F., Bourbonniere M., Tang D., Hu S.L., Weber J.M.;
RT "Nucleotide and deduced amino acid sequence of the bovine adenovirus
RT type 7 proteinase."
RL Nucleic Acids Res. 18:5567-5567(1990).
CC -!- FUNCTION: THIOL PROTEASE CLEAVING SPECIFIC GLY-ALA PEPTIDES IN A
CC NUMBER OF VIRAL PRECURSOR PROTEINS (TERMINAL, IIA, VI, VII, VIII,
CC IIA AND MU) WHICH ARE REQUIRED FOR VIRION MATURATION. ALSO CLEAVES
CC HOST CELLS CYTOSKELETAL KERATINS K7 AND K18.
CC -!- CATALYTIC ACTIVITY: Cleaves proteins of the adenovirus and its
CC host cell at two consensus sites: -Yaa-Xaa-Gly-Gly-|-Xaa- and
CC -Yaa-Xaa-Gly-Xaa-|-Gly- (in which Yaa is Met, Ile or Leu, and Xaa
CC is any amino acid).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53989; CAA37935.1; -.
CC PIR; S11434; S11434.
CC HSSP; P03252; IAVP.
CC MEROPS; C05.001; -.
CC InterPro; IPR000855; Peptidase_C5.
CC Pfam; PF00770; Peptidase_C5; 1.
CC PRINTS; PR00703; ADVENDOPTASE.
CC ProDom; PD003705; Peptidase_C5; 1.
CC Hydrolase; Thiol protease; Late protein.
FT ACT_SITE 55 55 BY SIMILARITY.
FT ACT_SITE 72 72 BY SIMILARITY.
FT ACT_SITE 122 122 BY SIMILARITY.
SQ SEQUENCE 202 AA; 23258 MW; 644113E7AEC2AC2 CRC64;
```



```

Query Match      19.1%  Score 62.5;  DB 1;  Length 1964;
Best Local Similarity 29.9%;  Pred. No. 24;
Matches 20;  Conservative 7;  Mismatches 11;  Indels 29;  Gaps 4;

QY      10  STKPGCSPIILRCAMLN-----PNNRLKD---TDCPGIKKCCGSC 49
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB      1101  SPKPGSPPL-----CACLSGFGGDCPLTPAPPAGCGPSPCLHNGTCTEPL-----GNP 1151

```

```

Qy 50 GMACTVP 56
Db 1152 GFOQTC 1158

RESULT 37
MT2A_MYTED STANDARD; PRT; 71 AA.
ID MT2A_MYTED
AC P80253; Q9GUJ1;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein 20-III isoform A (MT-20-IIIA) (MT-I).
OS Mytilus edulis (Blue mussel), and
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550, 29158;
[1]_TaxID=6550, 29158;
RN R1
RP SEQUENCE.
RC SPECIES=M.edulis;
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
RN R2
RP SEQUENCE FROM N.A.
RC SPECIES=M.galloprovincialis;
RA Ceratto N., Dondero F., Burlando B., Van de Loo J.W.H.P.,
RA Viarengo A.;
RT "Metallothionein genomic DNA of the mussel Mytilus
galloprovincialis.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC !- SUBUNIT: HOMODIMER.
CC !- INDUCTION: BY CADMIUM.
CC !- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF199020; AAC28538.1; -.
CC PIR; S39422; S39422.
CC InterPro; IPR003019; Metallothion.
CC InterPro; IPR001008; Mtlthion_misc.
CC Pfam; PF00131; metalthio; 1.
CC PRINTS; PR00875; MTWOLLUC.
CC Metal-binding; Metal-thiolate cluster; Cadmium.
CC INIT MET 0 0 BY SIMILARITY.
CC SEQUENCE 71 AA; 6899 MW; D4BFD9E599857657 CRC64;

Query Match 19.0%; Score 62; DB 1; Length 71;
Best Local Similarity 39.0%; Pred. No. 1.8;
Matches 16; Conservative 2; Mismatches 11; Indels 12; Gaps 3;

Qy 13 PGSCPIILIRCAMLPNRRCLKDTPCGIKKCCGCGMAC 53
Db 1 PGPCNCI-----ETNVICIGTCSG--KCCQ--CGDAC 29

RESULT 38
MT2B_MYTED STANDARD; PRT; 71 AA.
ID MT2B_MYTED
AC P80258;

```


DR Pfam; PF00041; fn3; 3.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF01030; Recep_L domain; 2.
 DR PRINTS; PR00014; FNTYPEIII.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODom; PD000001; Suk_pkinase; 1.
 DR SMART; SM00060; FN3; 3.
 DR SMART; SM00261; FU; 1.
 DR SMART; SM00219; TyRK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 KW Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; ATP-binding; Phosphorylation; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 716
 FT CHAIN INSULIN-LIKE PEPTIDE RECEPTOR, ALPHA-
 FT CHAIN (POTENTIAL).
 FT PROPEP 717 720
 FT CHAIN 721 1363
 FT CHAIN INSULIN-LIKE PEPTIDE RECEPTOR, BETA-
 FT CHAIN (POTENTIAL).
 FT DOMAIN 721 928
 FT TRANSSEM 929 949
 FT DOMAIN 950 1363
 FT DOMAIN CYTOPLASMIC (POTENTIAL).
 FT NP BIND 1000 1008
 FT BINDING 1028 1028
 FT ACT SITE 1148 1148
 FT MOD_RES 1174 1174
 FT CARBOHYD 51 51
 FT CARBOHYD 97 97
 FT CARBOHYD 137 137
 FT CARBOHYD 278 278
 FT CARBOHYD 483 483
 FT CARBOHYD 599 599
 FT CARBOHYD 617 617
 FT CARBOHYD 665 665
 FT CARBOHYD 666 666
 FT CARBOHYD 711 711
 FT CARBOHYD 732 732
 FT CARBOHYD 736 736
 FT CARBOHYD 743 743
 FT CARBOHYD 816 816
 FT CARBOHYD 885 885
 FT CARBOHYD 898 898
 SQ SEQUENCE 1363 AA; 15410 MW; 238120B4EAB1ED65 CRC64;

Query Match 18.8%; Score 61.5; DB 1; Length 1363;
 Best Local Similarity 23.7%; Pred. No. 23;
 Matches 18; Conservative 7; Mismatches 8; Indels 43; Gaps 5;

QY 14 GSCPIILIRCAMLNPPN-----RCLKDTDCPGI-----KKG-----44
 Db 252 GEC---LIQC-----PPDTQYKORRCITEECPTNTSVKWLHHRKCIPECPSGYTTDIN 304
 QY 45 -----CEGSCGCMAC 53
 Db 305 NPRLCTECGQCPKSC 320

RESULT 41
 TFF2 HUMAN
 ID TFF2_HUMAN STANDARD; PRT; 129 AA.
 AC Q03403; Q15854;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trefol factor 2 precursor (Spasmolytic polypeptide) (SP).
 GN TFF2 OR SML1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Stomach;
 RX MEDLINE=90151615; PubMed=2303034;
 RA Tomasetto C., Rio M.C., Gautier C., Wolf C., Hareuveni M., Chambon P.,
 Lathe R.;
 RT "hSP, the domain-duplicated homolog of pS2 protein, is co-expressed
 RT with pS2 in stomach but not in breast carcinoma.";
 RL EMBO J. 9:407-414(1990).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=97224512; PubMed=9070946;
 RA Seib T., Blin N., Hilgert K., Seifert M., Theisinger B., Engel M.,
 Dooley S., Zang K.D., Welter C.;
 RA "The three human trefol genes TFF1, TFF2, and TFF3 are located
 RT within a region of 55 kb on chromosome 21q22.3.";
 RL Genomics 40:200-202(1997).
 [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20408883; PubMed=10950923;
 RA Berry A., Scott H.S., Kudoh J., Tallor I., Korostishevsky M.,
 Wattenhofer M., Guipponi M., Barras C., Rossier C., Shibuya K.,
 Wang J., Kawasaki K., Asakawa S., Minoshima S., Shimizu N.,
 Antonarakis S., Bonne-Tamir B.;
 RA "Refined localization of autosomal recessive nonsyndromic deafness
 RT DFNB10 locus using 34 novel microsatellite markers, genomic structure,
 RT and exclusion of six known genes in the region.";
 RL Genomics 68:22-29(2000).
 CC -!- FUNCTION: INHIBITS GASTROINTESTINAL MOTILITY AND GASTRIC ACID
 CC SECRETION. COULD FUNCTION AS A STRUCTURAL COMPONENT OF GASTRIC
 CC MUCUS. POSSIBLY BY STABILIZING GLYCOPROTEINS IN THE MUCUS GEL
 CC THROUGH INTERACTIONS WITH CARBOHYDRATE SIDE CHAINS (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: STOMACH.
 CC -!- SIMILARITY: CONTAINS 2 P-TYPE (TREFOL) DOMAINS.
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chromocancer/Genes/TFF2ID264.html".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X51698; CAA35995.1; -;
 CC EMBL; U47292; AAB05397.1; -;
 CC EMBL; U47289; AAB05397.1; JOINED.
 CC EMBL; U47290; AAB05397.1; JOINED.
 CC EMBL; U47291; AAB05397.1; JOINED.
 CC EMBL; AB038162; BAB13730.1; -;
 CC PIR; S12371; S12371.
 CC HSSP; P01359; 1PCP.
 CC Genew; HGNC:11756; TFF2.
 CC MIM; 182590; -;
 CC InterPro; IPR000519; P_trefol.
 CC Pfam; PF00088; trefol_2.
 CC PRINTS; PR00680; PTREFOL.
 CC SMART; SM00018; P; 2.
 CC PROSITE; PS00025; P_TREFOL; 2.
 KW Repeat; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 129
 FT DOMAIN 30 72
 FT DOMAIN 80 121
 FT DISULFID 29 127
 FT DISULFID 31 58
 FT DISULFID 42 57
 FT DISULFID 52 69
 FT DISULFID 81 107
 FT POTENTIAL.
 FT TREFOL FACTOR 2.
 FT P-TYPE 1.
 FT P-TYPE 2.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.

FT DISULFID 91 106 BY SIMILARITY.
 FT DISULFID 101 118 BY SIMILARITY.
 FT CONFLICT 1 19 MGRDQALLAALLVGLCA -> RHGTARRPAGSAPRPGA
 FT CONFLICT 122 122 K -> N (IN REF. 1).
 FT CONFLICT 129 AA; 14284 MW; 6A296A11AD9CC5D CRC64;
 SQ SEQUENCE 129 AA; 14284 MW; 6A296A11AD9CC5D CRC64;
 Query Match 18.7%; Score 61; DB 1; Length 129;
 Best Local Similarity 36.7%; Pred. No. 3.7;
 Matches 18; Conservative 6; Mismatches 9; Indels 16; Gaps 5;
 Qy 10 STKGSCPIILIRCAMLPPNRLCKDTPC--PGI--KCCCEGCGMACF 54
 Db 23 SEKSPSC-----QCSRLSPHNR-----TNCGPPGITSQCDFNGC---CF 59
 RESULT 42
 YMW2 CAEEL STANDARD; PRT; 1246 AA.
 ID YMW2 CAEEL STANDARD; PRT; 1246 AA.
 AC P34504; P34505; P34506; P90907;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein K04H4.2 in chromosome III.
 CN K04H4.2
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton L., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; may be
 CC produced by alternative splicing.
 CC -!- SIMILARITY: TO CHITIN-BINDING MOTIFS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z27078; CA881587.1; -;
 CC EMBL; Z27078; CA881588.1; -;
 DR EMBL; Z27078; CA881588.1; -;
 DR PIR; S40992; S40992.
 DR PIR; S40993; S40993.
 DR PIR; S40994; S40994.
 DR HSSP; P10968; 2CWG.
 DR WormPep; K04H4.2a; CE19967.
 DR WormPep; K04H4.2b; CE19968.
 DR InterPro; IPR002557; Chitin_bind_PeRA.
 DR InterPro; IPR002899; WR1/EB.

DR SMART; SM00494; ChtBD2; 1.
 DR SMART; SM00289; WR1; 14.
 KW Hypothetical protein; Alternative splicing.
 FT VARSPLIC 166 529
 FT ANCTSDDEDPPTTFKCYOGCKKLVQPSRLTAFTCTKQVH
 FT CRANEHCFGGCCPKTIELAVIKSVLTKDNEHTKTEK
 FT LIIGDEVTRVKKCDIDIIICPMSECVGICCKQPPKARC
 FT NGMLALSIPVHCSLSDCPIASRCYKGCPCPLSESADST
 FT SDVGETTPVILKEEIIISTATKWKVKDKTSGVSNKNKCL
 FT STQRCDLHTLCPDPFTCSLSKCKCLNKHCPDGTGVPETSCQ
 FT SASNHDHCPSSSHKCTLLNKEHFACCCYSPGLVVEGSVTAEV
 FT SSECPIGSVEVDPFRFTGSCRYSLQCPSPYFCNQRGQ -> Q
 FT (IN ISOFORM B).
 FT DSDNTNTNPSQPTTKTKNTTKNSKSKHKKPKKDDVP
 FT LSDPLLDNDPFGPGYGFPEHLNLDVLIRAOGDGECPA
 FT GLHCDTAINLCCPLLPLTDPKPKKTKERRKQKQDGNEM
 FT EASANFPDSDPARFSSYSCGCMG -> VG (IN
 FT ISOFORM B).
 SQ SEQUENCE 1246 AA; 130610 MW; 4FA1A17D3F9606C4 CRC64;
 Query Match 18.7%; Score 61; DB 1; Length 1246;
 Best Local Similarity 47.6%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 Qy 32 CLKDTDCPGIKKCEGCGMA 52
 Db 209 CTSDEDCPTTFKCYOGCKLA 229
 RESULT 43
 ID ITRI MOMCH STANDARD; PRT; 30 AA.
 AC P10294;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Trypsin inhibitor I (MCTI-1).
 OS Momordica charantia (Bitter melon) (Balsam pear).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
 OX NCBI_TaxID=3673;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=892991812; PubMed=2738047;
 RA Hara S., Makino J., Ikenaka T.;
 RT "Amino acid sequences and disulfide bridges of serine proteinase
 RL J. Biochem. 105:88-92(1989).
 CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
 CC INHIBITORS.
 DR PIR; JX0057; JX0057.
 DR HSSP; P01074; LBXJ.
 DR InterPro; IPR000737; Squash.
 DR PPRINTS; PR00299; squash; 1.
 DR PRODOM; PD003401; Squash; 1.
 DR SMART; SM00286; PTI; 1.
 DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
 KW Serine protease inhibitor.
 FT ACT_SITE 6 7 REACTIVE BOND.
 FT DISULFID 4 21 BY SIMILARITY.
 FT DISULFID 11 23 BY SIMILARITY.
 FT DISULFID 17 29 BY SIMILARITY.
 SQ SEQUENCE 30 AA; 3413 MW; D5E1344365E34E9A CRC64;
 Query Match 18.3%; Score 60; DB 1; Length 30;
 Best Local Similarity 38.9%; Pred. No. 1.4;
 Matches 14; Conservative 3; Mismatches 9; Indels 10; Gaps 2;
 Qy 16 CPILLIRCAMLPPNRLCKDTPCGIKKC-CEGSCG 50

```

Db 4 CPrIL-----KQCKRSDSCPGCICMAHGFCG 30

RESULT 44
MT21 MYTED
ID MT21 MYTED STANDARD; PRT; 71 AA.
AC P80257;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein 20-I isoforms A and B (MT-20-IA and MT-20-IB).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RT forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
CC -!- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- INDUCTION: BY CADMIUM.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.
DR PIR; S39420; S39420.
DR InterPro; IPR001019; Metallthion.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PR00875; MTMOLLUSC.
KW Metal-binding; Metal-thiolate cluster; Cadmium.
FT VARIANT 66 66 S -> A (IN MT-20-IB).
SQ SEQUENCE 71 AA; 6944 MW; 4FD275FE34E81654 CRC64;

Query Match 18.3%; Score 60; DB 1; Length 71;
Best Local Similarity 39.0%; Pred. No. 2.9;
Matches 16; Conservative 1; Mismatches 12; Indels 12; Gaps 3;

QY 13 PGSCPILIRCAMLNPPNRLKDTDCPGIKKCCGSGCMAC 53
Db 1 GPCNCI-----ETNVCICGTGCGS--KCCR--CGDAC 29

RESULT 45
MT22 MYTED
ID MT22 MYTED STANDARD; PRT; 71 AA.
AC P80252;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein 20-II (MT-20-II).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RT forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Digestive gland;
RX MEDLINE=99206055; PubMed=10190057;
RA Barsyte D., White K.N., Lovejoy D.A.;
RT "Cloning and characterization of metallothionein cDNAs in the mussel
RT Mytilus edulis L. digestive gland.";

Comp. Biochem. Physiol. 122C:287-296(1999).
-!- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- INDUCTION: BY CADMIUM.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ005456; CAA06553.1; -.
DR PIR; S39421; S39421.
DR InterPro; IPR003019; Metallthion.
DR Pfam; PF001008; Mtlthion_misc.
DR PRINTS; PR00875; MTMOLLUSC.
KW Metal-binding; Metal-thiolate cluster; Cadmium.
FT INIT_MET 0
SQ SEQUENCE 71 AA; 6887 MW; 4FD275FE39857654 CRC64;

Query Match 18.3%; Score 60; DB 1; Length 71;
Best Local Similarity 39.0%; Pred. No. 2.9;
Matches 16; Conservative 1; Mismatches 12; Indels 12; Gaps 3;

QY 13 PGSCPILIRCAMLNPPNRLKDTDCPGIKKCCGSGCMAC 53
Db 1 GPCNCI-----ETNVCICGTGCGS--KCCR--CGDAC 29

RESULT 46
ZAN_MOUSE
ID ZAN_MOUSE STANDARD; PRT; 5376 AA.
AC O88799; O08647;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zonadhesin precursor.
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98123114; PubMed=9452463;
RA Gao Z., Garbers D.L.;
RT "Species diversity in the structure of zonadhesin, a sperm-specific
RT membrane protein containing multiple cell adhesion molecule-like
RT domains.";
RL J. Biol. Chem. 273:3415-3421(1998).
RN [2]
RP SEQUENCE OF 4864-5376 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=97271566; PubMed=9126492;
RA Gao Z., Harumi T., Garbers D.L.;
RT "Chromosome localization of the mouse zonadhesin gene and the human
RT zonadhesin gene (ZAN)";
RL Genomics 41:119-122(1997).
CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC SIGNALING.
CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC APICAL REGION OF THE SPERM HEAD.
CC -!- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
CC -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
CC ZONA PELLUCIDA.

```

	FT	DOMAIN	4504	4623	VWFD 22 (PARTIAL).
	FT	DOMAIN	4624	4743	VWFD 23 (PARTIAL).
	FT	DOMAIN	4744	4863	VWFD 24 (PARTIAL).
	FT	DOMAIN	4864	5261	VWFD 25.
	FT	DOMAIN	5259	5295	EGF-LIKE.
	FT	DISULFID	5263	5274	BY SIMILARITY.
	FT	DISULFID	5268	5283	BY SIMILARITY.
	FT	DISULFID	5285	5294	BY SIMILARITY.
	FT	CARBOHYD	339	339	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	499	499	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1216	1216	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1239	1239	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1314	1314	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1314	1814	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1814	1908	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1908	1908	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	1933	1933	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2028	2028	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2111	2111	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2142	2142	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2332	2332	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2533	2533	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2575	2575	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2692	2692	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	2812	2812	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3052	3052	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3065	3065	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3144	3144	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3172	3172	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3288	3288	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3292	3292	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	3782	3782	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4005	4005	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4136	4136	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4243	4243	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4254	4254	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4335	4335	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4376	4376	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	4586	4586	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	5136	5136	N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	5252	5252	N-LINKED (GLCNAC. .) (POTENTIAL).
	SQ	SEQUENCE	5376 AA;	579908 MW;	0E44B77DF2A2620 CRC64;
		Query Match	18.3%;	Score 60;	DB 1; Length 5376;
		Best Local Similarity	39.0%;	Pred. No. 1e+02;	
		Matches 16;	Conservative 1;	Mismatches 12;	Indels 12;
Oy	23	CAMLNPN---RCLXDTCPG-----IKKCEGSCGMAC 53			
Dd	5146	CAMTCPCANTVYQRQM--TFCPASCAKFVTPKVCGPCVEGC 5184			
		RESULT 47			
		ID_MT_EISFO	STANDARD;	PRT;	75 AA.
AC	P81695;	O96822;			
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DE	15-JUN-2002	(Rel. 41, Last annotation update)			
DD	Cadmium-metallothionein (MT)	(Fragment)			
GN	MT.	Eisenia foetida (Common handring worm) (Common dung-worm).			
OS	Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaeni				
OC	Lumbricina; Lumbricidae; Eisenia.				
OX	NCBI_TaxId=6396;				
RN	[1]				
RP	SEQUENCE OF 1-41.				
RA	Gruber C., Dallinger R., Hunziker P., Gehrig P., Berger B.,				
RA	Stuerzenbaum S.;				
RT	"(Cd)-metallothionein from Eisenia foetida: evidence for				
RT	posttranslational processing to a functional one-domain protein				
RL	Submitted (FEB-1999) to the SWISS-PROT data bank.				
RN	[2]				
RP	SEQUENCE OF 2-75 FROM N.A.				

RA Sturzenbaum S.R., Guppy L., Morgan A.J., Kille P.;
RT "Metallothionein isolated from the common branding worm Eisenia foetida";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: POSTERIOR ALIMENTARY CANAL.
CC -!- INDUCTION: BY CADMIUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ236886; CAAL5423.1; --
DR InterPro; IPR001396; EchnMethion4.
DR PRINTS; PR00873; MTECHINOIDEA.
KW Metal-binding; Metal-thiolate cluster; Cadmium.
FT NON TER 1 1
FT CONFLICT 11 11 A -> Q (IN REF. 1).
SQ SEQUENCE 75 AA; 7411 MW; FDE2C30BA606494B CRC64;

Query Match 18.2%; Score 59.5; DB 1; Length 75;
Best Local Similarity 51.9%; Pred. No. 3.4;
Matches 14; Conservative 1; Mismatches 7; Indels 5; Gaps 2;

Qy 31 RCLKXDCPGIKK-CC-----RGSCGM 52
| | | | | | | | | | : | | | |
Db 23 RCLKXSECLPGCKKLCADAEKRGCGNA 49
| | | | | | | | | | : | | | |

RESULT 48
YC92_CABEEL STANDARD; PRT; 621 AA.
ID YC92_CABEEL STANDARD; PRT; 621 AA.
AC P5115;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical zinc metalloproteinase T04G9.2 (EC 3.4.24.-).
GN T04G9.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Minx P.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U41274; AAA82461.1; --
DR HSSP; P07584; 11AE.
DR MEROPS; M12 UPA;
DR WormRep; T04G9.2; CE04883.
DR InterPro; IPR001506; Astacin.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR003582; ShKT.
DR Pfam; PF01400; Zn MTpeptdase.
DR PRINTS; PR00480; ASTACIN.
DR

DR SMART; SM00254; ShKT; 3.
DR SMART; SM00235; ZmMc; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc.
FT METAL 255 255 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 256 256 BY SIMILARITY.
FT METAL 259 259 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 265 265 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 621 AA; 68584 MW; A98D3721C8ADE489 CRC64;

Query Match 18.2%; Score 59.5; DB 1; Length 621;
Best Local Similarity 27.7%; Pred. No. 20;
Matches 18; Conservative 7; Mismatches 27; Indels 13; Gaps 3;

Qy 2 QBPVKGPSV-----TKPGSPILLRCAMLNPPNRCCLKDTCPCIKKCCGSCGM 51
| | | | | | | | | | : | | | |
Db 463 QBPVQPIQPLPPVPLPPTPTTDCEDLRVCLVLVSRYCKISQNF--MKSYYCAKSCGF 520
| | | | | | | | | | : | | | |

Qy 52 ACQVFP 56
| | | | | | | | | | : | | | |
Db 521 -CFKP 524
| | | | | | | | | | : | | | |

RESULT 49
NTC4_HUMAN STANDARD; PRT; 2003 AA.
ID NTC4_HUMAN STANDARD; PRT; 2003 AA.
AC Q99456; Q00306; Q99940; Q99458; Q9H3S8; Q9UII9; Q9UIJ0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurogenic locus notch homolog protein 4 precursor (Notch 4)
DE (hNotch4).
GN NOTCH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND POLYMORPHISM OF POLY-LEU.
RC TISSUE=Placenta;
RX MEDLINE=97311416; PubMed=9168133;
RA Sugaya K., Sasanuma S.-I., Nohata J., Kimura T., Fukagawa T.,
RA Nakamura Y., Ando A., Inoko H., Ikemura T., Mita K.;
RT "Gene organization of human NOTCH4 and (CTG)n polymorphism in this
RT human counterpart gene of mouse proto-oncogene Int3";
RL Gene 189:235-244 (1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Bone marrow, and Heart;
RX MEDLINE=98360091; PubMed=9693032;
RA Li L., Huang G.M., Banta A.B., Deng Y., Smith T., Dong P.,
RA Friedman C., Chen L., Trask B.J., Spies T., Rowen L., Hood L.;
RT "Cloning, characterization, and the complete 56.8-kilobase DNA
RT sequence of the human NOTCH4 gene";
RL Genomics 51:45-58 (1998).
RN [3]
RP SEQUENCE OF 1-503 FROM N.A., AND VARIANTS GLN-117 AND GLN-317.
RA Miyagawa T., Tokunaga K., Hojho H.;
RT "Human notch4 gene variant";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP IDENTIFICATION OF LIGANDS.
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
RT "Human ligands of the Notch receptor";
RL Am. J. Pathol. 154:785-794 (1999).
CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
CC Upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and

apoptotic programs. May regulate branching morphogenesis in the developing vascular system (By similarity).

-! SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-terminal fragment N(EC) which are probably linked by disulfide bonds (By similarity).

-! SUBCELLULAR LOCATION: Type I membrane protein. Following proteolytical processing NICD is translocated to the nucleus.

-! ALTERNATIVE PRODUCTS: 3 isoforms, 1 (shown here), 2 and 3; may be produced by alternative splicing.

-! TISSUE SPECIFICITY: Highly expressed in the heart, moderately in the lung and placenta and at low levels in the liver, skeletal muscle, kidney, pancreas, spleen, lymph node, thymus, bone marrow and fetal liver. No expression was seen in adult brain or peripheral blood leukocytes.

-! PTM: Synthesized in the endoplasmic reticulum as an inactive form which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active, ligand-accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NICD) from the membrane (By similarity).

-! PTM: Phosphorylated (By similarity).

-! POLYMORPHISM: The poly-Leu region of NOTCH4 (in the signal peptide) is polymorphic and the number of Leu varies in the population (from 6 to 12).

-! SIMILARITY: BELONGS TO THE NOTCH FAMILY.

-! SIMILARITY: CONTAINS 28 EGF-LIKE DOMAINS.

-! SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.

-! SIMILARITY: CONTAINS 5 ANK REPEATS.

-! CAUTION: Ref.1 sequence differs from that shown due to frameshifts in position 1438 to 1463.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D63395; BAA09708.1; ALT_FRAME.
EMBL; D86566; BAA13116.1; -.
EMBL; U95299; AAC32288.1; -.
EMBL; U89335; AAC63097.1; -.
EMBL; AB023961; BAB20317.1; -.
EMBL; AB024520; BAA88951.1; -.
EMBL; AB024578; BAA88952.1; -.
HSSP; P08709; 1BF9.
Genew; HGNC:7884; NOTCH4.
MIM; 164951; -.
InterPro; IPR002110; ANK.
InterPro; IPR000152; Asx hydroxyl.
InterPro; IPR000561; EGF-like.
InterPro; IPR000742; EGF 2.
InterPro; IPR001881; EGF_Ca.
InterPro; IPR001438; EGF_II.
InterPro; IPR000800; Notch.
Pfam; PF00008; EGF; 26.
Pfam; PF00023; ank; 6.
Pfam; PF00066; notch; 2.
PRINTS; PR00010; EGFBL00D.
PRINTS; PR00011; EGFAMININ.
PRINTS; PR00012; FNTYPEI.
SMART; SM00248; ANK; 5.
SMART; SM00179; EGF_Ca; 11.
SMART; SM00001; EGF_like; 15.
SMART; SM00004; NL; 2.
PROSITE; PS50088; ANK_REPEAT; 5.
PROSITE; PS50297; ANK_REPEAT_REGION; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; 11.
DR PROSITE; PS00022; EGF_1; 21.
DR PROSITE; PS01186; EGF_2; 28.
DR PROSITE; PS01187; EGF_CA; 9.
KW Receptor; Transcription regulation; Activator; Differentiation; Developmental protein; Repeat; ANK repeat; EGF-like domain; Transmembrane; Glycoprotein; Signal; Phosphorylation; Polymorphism; Triplet repeat expansion; Alternative splicing.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 2003 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 4.
FT CHAIN 1432 2003 NOTCH EXTRACELLULAR TRUNCATION (BY SIMILARITY).
FT CHAIN 1467 2003 NOTCH INTRACELLULAR DOMAIN (BY SIMILARITY).
FT DOMAIN 24 1447 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1448 1468 POTENTIAL.
FT DOMAIN 1469 2003 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 24 63 EGF-LIKE 1.
FT DOMAIN 64 115 EGF-LIKE 2.
FT DOMAIN 118 155 EGF-LIKE 3.
FT DOMAIN 156 192 EGF-LIKE 4.
FT DOMAIN 194 232 EGF-LIKE 5.
FT DOMAIN 234 274 EGF-LIKE 6.
FT DOMAIN 276 312 EGF-LIKE 7.
FT DOMAIN 314 353 EGF-LIKE 8.
FT DOMAIN 355 391 EGF-LIKE 9.
FT DOMAIN 392 430 EGF-LIKE 10.
FT DOMAIN 432 473 EGF-LIKE 11.
FT DOMAIN 475 511 EGF-LIKE 12.
FT DOMAIN 513 549 EGF-LIKE 13.
FT DOMAIN 551 587 EGF-LIKE 14.
FT DOMAIN 589 625 EGF-LIKE 15.
FT DOMAIN 626 659 EGF-LIKE 16.
FT DOMAIN 661 689 EGF-LIKE 17.
FT DOMAIN 691 727 EGF-LIKE 18.
FT DOMAIN 729 765 EGF-LIKE 19.
FT DOMAIN 767 803 EGF-LIKE 20.
FT DOMAIN 806 842 EGF-LIKE 21.
FT DOMAIN 844 880 EGF-LIKE 22.
FT DOMAIN 882 928 EGF-LIKE 23.
FT DOMAIN 930 966 EGF-LIKE 24.
FT DOMAIN 968 1004 EGF-LIKE 25.
FT DOMAIN 1006 1044 EGF-LIKE 26.
FT DOMAIN 1046 1085 EGF-LIKE 27.
FT DOMAIN 1087 1126 EGF-LIKE 28.
FT DOMAIN 1130 1171 EGF-LIKE 29.
FT DOMAIN 1171 1212 POLY-ARG.
FT REPEAT 1165 1212 LIN/NOTCH 1.
FT REPEAT 1213 1246 LIN/NOTCH 2.
FT REPEAT 1247 1286 LIN/NOTCH 3.
FT REPEAT 1633 1665 ANK 1.
FT REPEAT 1666 1698 ANK 2.
FT REPEAT 1700 1732 ANK 3.
FT REPEAT 1733 1765 ANK 4.
FT REPEAT 1766 1798 ANK 5.
FT DISULFID 28 41 BY SIMILARITY.
FT DISULFID 35 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 68 80 BY SIMILARITY.
FT DISULFID 74 103 BY SIMILARITY.
FT DISULFID 105 114 BY SIMILARITY.
FT DISULFID 122 133 BY SIMILARITY.
FT DISULFID 127 143 BY SIMILARITY.
FT DISULFID 145 154 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 198 211 BY SIMILARITY.
FT DISULFID 205 220 BY SIMILARITY.
FT DISULFID 222 231 BY SIMILARITY.
FT DISULFID 238 249 BY SIMILARITY.
FT DISULFID 243 262 BY SIMILARITY.
FT DISULFID 264 273 BY SIMILARITY.

FT DISULFID 168 BY SIMILARITY. 177
FT DISULFID 250 BY SIMILARITY. 262
FT DISULFID 257 BY SIMILARITY. 271
FT DISULFID 273 BY SIMILARITY. 286
FT DISULFID 292 BY SIMILARITY. 304
FT DISULFID 299 BY SIMILARITY. 313
FT DISULFID 315 BY SIMILARITY. 328
FT DISULFID 453 BY SIMILARITY. 465
FT DISULFID 460 BY SIMILARITY. 474
FT DISULFID 476 BY SIMILARITY. 488
FT DISULFID 494 BY SIMILARITY. 504
FT DISULFID 499 BY SIMILARITY. 513
FT DISULFID 515 BY SIMILARITY. 528
FT DISULFID 534 BY SIMILARITY. 546
FT DISULFID 541 BY SIMILARITY. 555
FT DISULFID 557 BY SIMILARITY. 570
FT DISULFID 576 BY SIMILARITY. 587
FT DISULFID 582 BY SIMILARITY. 596
FT DISULFID 598 BY SIMILARITY. 611
FT DISULFID 617 BY SIMILARITY. 628
FT DISULFID 623 BY SIMILARITY. 637
FT DISULFID 639 BY SIMILARITY. 652
FT DISULFID 727 BY SIMILARITY. 739
FT DISULFID 734 BY SIMILARITY. 748
FT DISULFID 750 BY SIMILARITY. 763
FT DISULFID 769 BY SIMILARITY. 781
FT DISULFID 776 BY SIMILARITY. 790
FT DISULFID 792 BY SIMILARITY. 805
FT DISULFID 811 BY SIMILARITY. 821
FT DISULFID 816 BY SIMILARITY. 830
FT DISULFID 832 BY SIMILARITY. 845
FT DISULFID 914 BY SIMILARITY. 926
FT DISULFID 921 BY SIMILARITY. 935
FT DISULFID 937 BY SIMILARITY. 950
FT DISULFID 1032 BY SIMILARITY. 1044
FT DISULFID 1039 BY SIMILARITY. 1053
FT DISULFID 1055 BY SIMILARITY. 1068
FT DISULFID 1074 BY SIMILARITY. 1086
FT DISULFID 1081 BY SIMILARITY. 1095
FT DISULFID 1097 BY SIMILARITY. 1111
FT DISULFID 1117 BY SIMILARITY. 1129
FT DISULFID 1124 BY SIMILARITY. 1138
FT DISULFID 1140 BY SIMILARITY. 1153
FT DISULFID 1159 BY SIMILARITY. 1171
FT DISULFID 1166 BY SIMILARITY. 1180
FT DISULFID 1182 BY SIMILARITY. 1195
FT DISULFID 1201 BY SIMILARITY. 1212
FT DISULFID 1208 BY SIMILARITY. 1221
FT DISULFID 1223 BY SIMILARITY. 1236
FT DISULFID 1242 BY SIMILARITY. 1254
FT DISULFID 1249 BY SIMILARITY. 1263
FT DISULFID 1265 BY SIMILARITY. 1278
FT DISULFID 1284 BY SIMILARITY. 1296
FT DISULFID 1291 BY SIMILARITY. 1305
FT DISULFID 1307 BY SIMILARITY. 1320
FT DISULFID 1326 BY SIMILARITY. 1339
FT DISULFID 1333 BY SIMILARITY. 1348
FT DISULFID 1350 BY SIMILARITY. 1361
FT DISULFID 1367 BY SIMILARITY. 1380
FT DISULFID 1374 BY SIMILARITY. 1389
FT DISULFID 1391 BY SIMILARITY. 1402
FT DISULFID 1408 BY SIMILARITY. 1420
FT DISULFID 1415 BY SIMILARITY. 1429
FT DISULFID 1431 BY SIMILARITY. 1444
FT DISULFID 1450 BY SIMILARITY. 1461
FT DISULFID 1456 BY SIMILARITY. 1470
FT DISULFID 1472 BY SIMILARITY. 1485
FT DISULFID 1491 BY SIMILARITY. 1502
FT DISULFID 1497 BY SIMILARITY. 1511
FT DISULFID 1513 BY SIMILARITY. 1526
FT DISULFID 1610 BY SIMILARITY. 1622
FT DISULFID 1617 BY SIMILARITY. 1631
FT DISULFID 1633 BY SIMILARITY. 1646

FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.
Query Match 18.0%; Score 59; DB 1; Length 2871;
Best Local Similarity 27.5%; Pred. No. 78; Indels 16; Gaps 4;
Matches 19; Conservative 8; Mismatches 26; Indels 16; Gaps 4;
QY 3 BPVKGVPSTKPGSCPII--LIRCAMLPNPPNRCLKDTDCFGIKKCE-----GSC----- 49
Db 1976 ECVLDPGKCAPGTQNLDSYRC--ICPPGYSLQNDKCEDICVEEPEICALGTCSNTE 2033
QY 50 -GMACFVPQ 57
Db 2034 GSFKCLCPE 2042
RESULT 51
TENX HUMAN
ID TENX HUMAN STANDARD; PRT: 4289 AA.
AC P22105; P78530; P78531; Q08424; Q9UMG7;
DT 01-AUG-1991 (Rel. 19, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tenascin-X precursor (TN-X) (Hexabrachion-like).
GN TNXB OR TNX OR XB OR HXBL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA Banta A., Schwartzell S., Smith T.M., Spies T., Hood L.;
RT "Sequence determination of 300 kilobases of the human class III MHC
locus".
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-747 AND 1687-1944 FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=93300909; PubMed=7686164;
RA Bristow J., Tee M.K., Gitelman S.E., Mellon S.H., Miller W.L.;
RT "Tenascin-X: a novel extracellular matrix protein encoded by the human
J. Cell Biol. 122:265-278(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM XB-SHORT).
RC TISSUE=Adrenal gland;
RX MEDLINE=96015044; PubMed=8530023;
RA Tee M.K., Thomson A.A., Bristow J., Miller W.L.;
RT "Sequences promoting the transcription of the human XA gene
overlapping P450c21A correctly predict the presence of a novel,
adrenal-specific, truncated form of tenascin-X".
RL Genomics 28:171-178(1995).
RN [4]
RP SEQUENCE OF 1-23 FROM N.A.
RC TISSUE=Petal adrenal gland;
RX MEDLINE=97081760; PubMed=8923003;
RA Speck M., Barry F., Miller W.L.;
RT "Alternate promoters and alternate splicing of human tenascin-X, a
gene with 5' and 3' ends buried in other genes".
RL Hum. Mol. Genet. 5:1749-1758(1996).
RN [5]
RP SEQUENCE OF 3470-4289 FROM N.A.
RX MEDLINE=89367293; PubMed=2475872;
RA Morel Y., Bristow J., Gitelman S.E., Miller W.L.;
RT "Transcript encoded on the opposite strand of the human steroid 21-
hydroxylase/complement component C4 gene locus".
RL Proc. Natl. Acad. Sci. U.S.A. 86:6582-6586(1989).
RN [6]
RP DISEASE.
RX PubMed=11642233;
RA Schalkwijk J., Zweers M.C., Steijlen P.M., Dean W.B., Taylor G.,

FT	DOMAIN	369	EGF-LIKE 8.
FT	DOMAIN	400	EGF-LIKE 9.
FT	DOMAIN	431	EGF-LIKE 10.
FT	DOMAIN	462	EGF-LIKE 11.
FT	DOMAIN	493	EGF-LIKE 12.
FT	DOMAIN	524	EGF-LIKE 13.
FT	DOMAIN	555	EGF-LIKE 14.
FT	DOMAIN	586	EGF-LIKE 15.
FT	DOMAIN	617	EGF-LIKE 16.
FT	DOMAIN	648	EGF-LIKE 17.
FT	DOMAIN	684	EGF-LIKE 18.
FT	DOMAIN	715	EGF-LIKE 19.
FT	DOMAIN	792	FIBRONECTIN TYPE-III 1.
FT	DOMAIN	901	COILED COIL (POTENTIAL).
FT	DOMAIN	941	FIBRONECTIN TYPE-III 2.
FT	DOMAIN	1047	FIBRONECTIN TYPE-III 3.
FT	DOMAIN	1149	FIBRONECTIN TYPE-III 4.
FT	DOMAIN	1246	FIBRONECTIN TYPE-III 5.
FT	DOMAIN	1348	FIBRONECTIN TYPE-III 6.
FT	DOMAIN	1459	FIBRONECTIN TYPE-III 7.
FT	DOMAIN	1561	FIBRONECTIN TYPE-III 8.
FT	DOMAIN	1659	FIBRONECTIN TYPE-III 9.
FT	DOMAIN	1756	FIBRONECTIN TYPE-III 10.
FT	DOMAIN	1856	FIBRONECTIN TYPE-III 11.
FT	DOMAIN	1962	FIBRONECTIN TYPE-III 12.
FT	DOMAIN	2069	FIBRONECTIN TYPE-III 13.
FT	DOMAIN	2167	FIBRONECTIN TYPE-III 14.
FT	DOMAIN	2266	FIBRONECTIN TYPE-III 15.
FT	DOMAIN	2365	FIBRONECTIN TYPE-III 16.
FT	DOMAIN	2471	FIBRONECTIN TYPE-III 17.
FT	DOMAIN	2582	FIBRONECTIN TYPE-III 18.
FT	DOMAIN	2688	FIBRONECTIN TYPE-III 19.
FT	DOMAIN	2794	FIBRONECTIN TYPE-III 20.
FT	DOMAIN	2889	FIBRONECTIN TYPE-III 21.
FT	DOMAIN	2997	FIBRONECTIN TYPE-III 22.
FT	DOMAIN	3105	FIBRONECTIN TYPE-III 23.
FT	DOMAIN	3211	FIBRONECTIN TYPE-III 24.
FT	DOMAIN	3307	FIBRONECTIN TYPE-III 25.
FT	DOMAIN	3399	FIBRONECTIN TYPE-III 26.
FT	DOMAIN	3494	FIBRONECTIN TYPE-III 27.
FT	DOMAIN	3601	FIBRONECTIN TYPE-III 28.
FT	DOMAIN	3699	FIBRONECTIN TYPE-III 29.
FT	DOMAIN	3801	FIBRONECTIN TYPE-III 30.
FT	DOMAIN	3890	FIBRONECTIN TYPE-III 31.
FT	DOMAIN	3978	FIBRONECTIN TYPE-III 32.
FT	DOMAIN	4071	FIBRONECTIN TYPE-III 33.
FT	DOMAIN	4289	FIBRONECTIN TYPE-III 34.
FT	DOMAIN	4718	FIBRONECTIN TYPE-III 35.
FT	DOMAIN	4750	FIBRONECTIN TYPE-III 36.
FT	DOMAIN	4879	FIBRONECTIN TYPE-III 37.
FT	DOMAIN	5022	FIBRONECTIN TYPE-III 38.
FT	DOMAIN	5128	FIBRONECTIN TYPE-III 39.
FT	DOMAIN	5234	FIBRONECTIN TYPE-III 40.
FT	DOMAIN	5340	FIBRONECTIN TYPE-III 41.
FT	DOMAIN	5446	FIBRONECTIN TYPE-III 42.
FT	DOMAIN	5552	FIBRONECTIN TYPE-III 43.
FT	DOMAIN	5658	FIBRONECTIN TYPE-III 44.
FT	DOMAIN	5764	FIBRONECTIN TYPE-III 45.
FT	DOMAIN	5870	FIBRONECTIN TYPE-III 46.
FT	DOMAIN	5976	FIBRONECTIN TYPE-III 47.
FT	DOMAIN	6082	FIBRONECTIN TYPE-III 48.
FT	DOMAIN	6188	FIBRONECTIN TYPE-III 49.
FT	DOMAIN	6294	FIBRONECTIN TYPE-III 50.
FT	DOMAIN	6400	FIBRONECTIN TYPE-III 51.
FT	DOMAIN	6506	FIBRONECTIN TYPE-III 52.
FT	DOMAIN	6612	FIBRONECTIN TYPE-III 53.
FT	DOMAIN	6718	FIBRONECTIN TYPE-III 54.
FT	DOMAIN	6824	FIBRONECTIN TYPE-III 55.
FT	DOMAIN	6930	FIBRONECTIN TYPE-III 56.
FT	DOMAIN	7036	FIBRONECTIN TYPE-III 57.
FT	DOMAIN	7142	FIBRONECTIN TYPE-III 58.
FT	DOMAIN	7248	FIBRONECTIN TYPE-III 59.
FT	DOMAIN	7354	FIBRONECTIN TYPE-III 60.
FT	DOMAIN	7460	FIBRONECTIN TYPE-III 61.
FT	DOMAIN	7566	FIBRONECTIN TYPE-III 62.
FT	DOMAIN	7672	FIBRONECTIN TYPE-III 63.
FT	DOMAIN	7778	FIBRONECTIN TYPE-III 64.
FT	DOMAIN	7884	FIBRONECTIN TYPE-III 65.
FT	DOMAIN	7990	FIBRONECTIN TYPE-III 66.
FT	DOMAIN	8096	FIBRONECTIN TYPE-III 67.
FT	DOMAIN	8202	FIBRONECTIN TYPE-III 68.
FT	DOMAIN	8308	FIBRONECTIN TYPE-III 69.
FT	DOMAIN	8414	FIBRONECTIN TYPE-III 70.
FT	DOMAIN	8520	FIBRONECTIN TYPE-III 71.
FT	DOMAIN	8626	FIBRONECTIN TYPE-III 72.
FT	DOMAIN	8732	FIBRONECTIN TYPE-III 73.
FT	DOMAIN	8838	FIBRONECTIN TYPE-III 74.
FT	DOMAIN	8944	FIBRONECTIN TYPE-III 75.
FT	DOMAIN	9050	FIBRONECTIN TYPE-III 76.
FT	DOMAIN	9156	FIBRONECTIN TYPE-III 77.
FT	DOMAIN	9262	FIBRONECTIN TYPE-III 78.
FT	DOMAIN	9368	FIBRONECTIN TYPE-III 79.
FT	DOMAIN	9474	FIBRONECTIN TYPE-III 80.
FT	DOMAIN	9580	FIBRONECTIN TYPE-III 81.
FT	DOMAIN	9686	FIBRONECTIN TYPE-III 82.
FT	DOMAIN	9792	FIBRONECTIN TYPE-III 83.
FT	DOMAIN	9898	FIBRONECTIN TYPE-III 84.
FT	DOMAIN	10004	FIBRONECTIN TYPE-III 85.
FT	DOMAIN	10110	FIBRONECTIN TYPE-III 86.
FT			

16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Sorting-related receptor precursor (Sorting protein-related receptor
containing LDLR class A repeats) (SorLA-1) (Low-density,
LDL protein receptor relative with 11 ligand-binding repeats) (LDLR
relative with 11 ligand-binding repeats) (LR11).
SOL1.
Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC TISSUE=Brain;
RX MEDLINE=97301565; PubMed=9157966;
RA Nykjaer L., Yamazaki H., Bujo H., Kusunoki J., Kanaki T., Seimiya K.,
Moriyasaki N., Nimf J., Schneider W.J., Saito Y.;
RT "A novel mosaic protein containing LDL receptor elements is highly
conserved in humans and chickens.";
RL Aterioscler. Thromb. Vasc. Biol. 17:996-1002(1997).
[2]
SEQUENCE FROM N.A.; SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.
RP TISSUE=Brain, and T-cell;
RC TISSUE=Brain, and T-cell;
RX MEDLINE=97094912; PubMed=8940146;
RA Jacobsen L., Madsen P., Moestrup S.K., Lund A.H., Tommerup N.,
Nykjaer A., Sottrup-Jensen L., Gliemann J., Petersen C.M.;
RT "Molecular characterization of a novel human hybrid-type receptor that
binds the alpha2-macroglobulin receptor-associated protein.";
RL J. Biol. Chem. 271:31379-31383(1996).
CC -1- FUNCTION: LIKELY TO BE A MULTIFUNCTIONAL ENDOCYTIC RECEPTOR, THAT
MAY BE IMPLICATED IN THE UPTAKE OF LIPOPROTEINS AND OF PROTEASES.
CC BINDS LDL, THE MAJOR CHOLESTEROL-CARRYING LIPOPROTEIN OF PLASMA,
CC AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS. BINDS THE RECEPTOR-
CC ASSOCIATED PROTEIN (RAP). COULD PLAY A ROLE IN CELL-CELL
CC INTERACTION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN BRAIN, WHERE IT IS MOST
CC ABUNDANT IN THE CEREBELLUM, CEREBRAL CORTEX AND THE OCCIPITAL
CC POLE; LOW EXPRESSION IN THE PUTAMEN AND THE THALAMUS, ACCORDING TO
CC REF.1, FOUND IN SPINAL CORD, TESTIS, LIVER, KIDNEY AND PANCREAS
CC WITH DETECTABLE LEVELS IN PLACENTA, LUNG AND HEART. ACCORDING TO
CC REF.2, EXPRESSED IN THE PROSTATE, OVARY, THYROID AND SPLEEN, BUT
CC NOT FOUND IN KIDNEY, LIVER, LUNG, SKELETAL MUSCLE, BONE MARROW AND
CC ADRENALS.
CC -1- PTM: THE PROPEPTIDE REMOVED IN THE N-TERMINUS MAY BE CLEAVED BY
CC FURIN OR HOMOLOGOUS PROTEASES.
CC -1- SIMILARITY: CONTAINS 5 BNR REPEATS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 11 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y08110; CAAG9325.1; -;
CC EMBL; U60975; AAC50891.2; -;
CC HSSP; P01130; IAJJ.
CC Genew; HGNC:11185; SORL1.
CC MIM; 602005; -;
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR003962; FNIII repeat.
CC InterPro; IPR002860; GH BNR.
CC InterPro; IPR002172; LDL_recept.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00041; fn3; 5.
CC Pfam; PF00057; ldl_recept_a; 11.

RC STRAIN=ATCC VR-896;
RA Davison A.J., Wright K.M., Harrach B.;
RT "Phylogenetic position of an amphibian adenovirus";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIOLE PROTEASE CLEAVING SPECIFIC GLY-ALA PEPTIDES IN A
CC NUMBER OF VIRAL PRECURSOR PROTEINS (TERMINAL, IIA, VI, VII, VIII,
CC IIA AND MU) WHICH ARE REQUIRED FOR VIRION MATURATION. ALSO CLEAVES
CC HOST CELLS CYTOSKELETAL KERATINS K7 AND K18.
CC -!- CATALYTIC ACTIVITY: Cleaves proteins of the adenovirus and its
CC host cell at two consensus sites: -Yaa-Xaa-Gly-Gly-|-Xaa- and
CC -Yaa-Xaa-Gly-Xaa-|-Gly- (in which Yaa is Met, Ile or Leu, and Xaa
CC is any amino acid).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C5.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF224336; AAF86933.1; -
CC HSSP; P03252; 1AVP.
CC MEROPS; C05.001; -
CC InterPro: IPR000855; Peptidase C5.
CC Pfam; PF00770; Peptidase C5; 1.
CC PRINTS; PR00703; ADENOPTASE.
CC ProDom; PD003705; Peptidase C5; 1.
CC HydroLase; Thiol protease; Late protein.
CC ACT_SITE 54 BY SIMILARITY.
CC ACT_SITE 71 BY SIMILARITY.
CC ACT_SITE 121 BY SIMILARITY.
CC ACT_SITE 121 BY SIMILARITY.
CC SEQUENCE 204 AA; 23222 MW; 91DC0B01544577A5 CRC64;

Query Match 17.6%; Score 57.5; DB 1; Length 204;
Best Local Similarity 28.6%; Pred. No. 13;
Matches 10; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 20 LIRCAMLNPPNRLKDTCPGFKKC-CRSGSCGMAC 53
DB 91 ILKSTALTKPSCIKLVKSKEAVQCTSAACGLFC 125

RESULT 58
LRP2_RAT
ID LRP2_RAT STANDARD; PRT; 4660 AA.
AC P98158;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
DE (Glycoprotein 330).
DE LRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=95024033; PubMed=7937880;
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;
RT "Complete cloning and sequencing of rat gp330/megalin, a
RT distinctive member of the low density lipoprotein receptor gene
RT family";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
RL [2]
RP FUNCTION.
RX MEDLINE=95386696; PubMed=7544804;
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,
RA Norris K., Gilemann J., Christensen E.I.;
RT "Evidence that epithelial glycoprotein 330/megalin mediates uptake of

polybasic drugs.";
J. Clin. Invest. 96:1404-1413(1995).
[3]
TT TISSUE SPECIFICITY.
RX MEDLINE=94172242; PubMed=7510321;
RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
RA Andres G., McCluskey R.T.;
RT "Organ distribution in rats of two members of the low-density
RT lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the
RT receptor-associated protein (RAP).";
RL J. Histochem. Cytochem. 42:531-542(1994).
CC -!- FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,
CC PLASMINOGEN ACTIVATOR-PLASMINOGEN ACTIVATOR INHIBITOR TYPE I
CC COMPLEX, APOLOPROTEIN E-ENRICHED BETA-VLDL, LIPOPROTEIN LIPASE,
CC LACTOFERRIN, CLUSTERIN AND CALCIUM.
CC -!- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS
CC APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B.
CC -!- SUBUNIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-
CC ASSOCIATED PROTEIN (RAP).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXPRESSED IN
CC CLATHRIN-COATED PITS; A SOLUBLE FORM IS POSSIBLY DERIVED BY
CC CLEAVAGE AT THE CELL SURFACE.
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF KIDNEY GLOMERULUS AND
CC PROXIMAL TUBULE, LUNG, EPIDIDYMIS, YOLK SAC, AMONG OTHERS.
CC -!- SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS.
CC -!- SIMILARITY: CONTAINS 17 LDL-RECEPTOR CLASS B DOMAINS.
CC -!- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L34049; AAA51369.1; -
CC HSSP; Q07954; 1CR8.
CC GlycoSuiteDB; P98158; -
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001881; EGF_Ca.
CC InterPro: IPR002172; LDL_receptor_A.
CC InterPro: IPR000033; Ldl_receptor_rep.
CC Pfam; PF00057; ldl_recept_a; 36.
CC Pfam; PF00058; ldl_recept_b; 33.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 3.
CC SMART; SM00001; EGF_Like; 15.
CC SMART; SM00192; LDLA; 36.
CC SMART; SM00135; LY; 35.
CC PROSITE; PS00010; ASX_HYDROXYL; 4.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 8.
CC PROSITE; PS01187; EGF_CA; 3.
CC PROSITE; PS01209; LDLRA_1; 31.
CC PROSITE; PS00068; LDLRA_2; 36.
CC Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
CC Receptor; EGF-like domain; SH3-binding; Signal.
FT SIGNAL 1 25
FT CHAIN 26 4660
FT
FT LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT PROTEIN 2.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT LDL-RECEPTOR CLASS A 1.
FT LDL-RECEPTOR CLASS A 2.
FT LDL-RECEPTOR CLASS A 3.
FT LDL-RECEPTOR CLASS A 4.
FT LDL-RECEPTOR CLASS A 5.
FT LDL-RECEPTOR CLASS A 6.
FT LDL-RECEPTOR CLASS A 7.
FT EGF-LIKE 1.
FT
FT DOMAIN 26 4425
FT DOMAIN 4426 4446
FT TRANSMEM 4447 4660
FT DOMAIN 4447 4660
FT DOMAIN 26 64
FT DOMAIN 65 105
FT DOMAIN 106 144
FT DOMAIN 145 181
FT DOMAIN 181 219
FT DOMAIN 220 258
FT DOMAIN 263 307
FT DOMAIN 308 346

FT DOMAIN 347 385
 FT DOMAIN 435 476
 FT DOMAIN 478 519
 FT DOMAIN 521 566
 FT DOMAIN 568 611
 FT DOMAIN 612 652
 FT DOMAIN 658 704
 FT DOMAIN 752 793
 FT DOMAIN 795 835
 FT DOMAIN 837 879
 FT DOMAIN 881 923
 FT DOMAIN 969 1013
 FT DOMAIN 1023 1061
 FT DOMAIN 1064 1103
 FT DOMAIN 1108 1146
 FT DOMAIN 1148 1186
 FT DOMAIN 1186 1225
 FT DOMAIN 1229 1269
 FT DOMAIN 1270 1308
 FT DOMAIN 1311 1351
 FT DOMAIN 1350 1390
 FT DOMAIN 1391 1430
 FT DOMAIN 1479 1520
 FT DOMAIN 1522 1563
 FT DOMAIN 1567 1609
 FT DOMAIN 1611 1654
 FT DOMAIN 1656 1696
 FT DOMAIN 1701 1742
 FT DOMAIN 1791 1832
 FT DOMAIN 1834 1882
 FT DOMAIN 1884 1930
 FT DOMAIN 1932 1972
 FT DOMAIN 1973 2013
 FT DOMAIN 2019 2060
 FT DOMAIN 2108 2156
 FT DOMAIN 2158 2201
 FT DOMAIN 2203 2245
 FT DOMAIN 2247 2289
 FT DOMAIN 2291 2332
 FT DOMAIN 2332 2384
 FT DOMAIN 2432 2477
 FT DOMAIN 2479 2518
 FT DOMAIN 2520 2562
 FT DOMAIN 2564 2604
 FT DOMAIN 2605 2647
 FT DOMAIN 2652 2694
 FT DOMAIN 2699 2739
 FT DOMAIN 2740 2778
 FT DOMAIN 2779 2820
 FT DOMAIN 2821 2862
 FT DOMAIN 2863 2903
 FT DOMAIN 2906 2947
 FT DOMAIN 2948 2992
 FT DOMAIN 2993 3031
 FT DOMAIN 3032 3072
 FT DOMAIN 3075 3112
 FT DOMAIN 3113 3153
 FT DOMAIN 3154 3194
 FT DOMAIN 3241 3282
 FT DOMAIN 3284 3333
 FT DOMAIN 3335 3377
 FT DOMAIN 3379 3420
 FT DOMAIN 3421 3461
 FT DOMAIN 3467 3511
 FT DOMAIN 3512 3552
 FT DOMAIN 3553 3593
 FT DOMAIN 3594 3634
 FT DOMAIN 3635 3675
 FT DOMAIN 3678 3718
 FT DOMAIN 3719 3758
 FT DOMAIN 3759 3797
 FT DOMAIN 3798 3836
 FT DOMAIN 3842 3882

EGF-LIKE 2.
 LDL-RECEPTOR CLASS B 1.
 LDL-RECEPTOR CLASS B 2.
 LDL-RECEPTOR CLASS B 3.
 LDL-RECEPTOR CLASS B 4.
 LDL-RECEPTOR CLASS B 5.
 EGF-LIKE 3.
 LDL-RECEPTOR CLASS B 6.
 LDL-RECEPTOR CLASS B 7.
 LDL-RECEPTOR CLASS B 8.
 LDL-RECEPTOR CLASS B 9.
 EGF-LIKE 4.
 LDL-RECEPTOR CLASS A 8.
 LDL-RECEPTOR CLASS A 9.
 LDL-RECEPTOR CLASS A 10.
 LDL-RECEPTOR CLASS A 11.
 LDL-RECEPTOR CLASS A 12.
 LDL-RECEPTOR CLASS A 13.
 LDL-RECEPTOR CLASS A 14.
 LDL-RECEPTOR CLASS A 15.
 EGF-LIKE 5.
 EGF-LIKE 6.
 LDL-RECEPTOR CLASS B 10.
 LDL-RECEPTOR CLASS B 11.
 LDL-RECEPTOR CLASS B 12.
 LDL-RECEPTOR CLASS B 13.
 LDL-RECEPTOR CLASS B 14.
 EGF-LIKE 7.
 LDL-RECEPTOR CLASS B 15.
 LDL-RECEPTOR CLASS B 16.
 LDL-RECEPTOR CLASS B 17.
 LDL-RECEPTOR CLASS B 18.
 LDL-RECEPTOR CLASS B 19.
 EGF-LIKE 8.
 LDL-RECEPTOR CLASS B 20.
 LDL-RECEPTOR CLASS B 21.
 LDL-RECEPTOR CLASS B 22.
 LDL-RECEPTOR CLASS B 23.
 LDL-RECEPTOR CLASS B 24.
 EGF-LIKE 9.
 LDL-RECEPTOR CLASS B 25.
 LDL-RECEPTOR CLASS B 26.
 LDL-RECEPTOR CLASS B 27.
 LDL-RECEPTOR CLASS B 28.
 LDL-RECEPTOR CLASS B 29.
 EGF-LIKE 10.
 LDL-RECEPTOR CLASS A 16.
 LDL-RECEPTOR CLASS A 17.
 LDL-RECEPTOR CLASS A 18.
 LDL-RECEPTOR CLASS A 19.
 LDL-RECEPTOR CLASS A 20.
 LDL-RECEPTOR CLASS A 21.
 LDL-RECEPTOR CLASS A 22.
 LDL-RECEPTOR CLASS A 23.
 LDL-RECEPTOR CLASS A 24.
 LDL-RECEPTOR CLASS A 25.
 EGF-LIKE 11.
 EGF-LIKE 12.
 LDL-RECEPTOR CLASS B 30.
 LDL-RECEPTOR CLASS B 31.
 LDL-RECEPTOR CLASS B 32.
 LDL-RECEPTOR CLASS B 33.
 LDL-RECEPTOR CLASS B 34.
 EGF-LIKE 13.
 LDL-RECEPTOR CLASS A 26.
 LDL-RECEPTOR CLASS A 27.
 LDL-RECEPTOR CLASS A 28.
 LDL-RECEPTOR CLASS A 29.
 LDL-RECEPTOR CLASS A 30.
 LDL-RECEPTOR CLASS A 31.
 LDL-RECEPTOR CLASS A 32.
 LDL-RECEPTOR CLASS A 33.
 LDL-RECEPTOR CLASS A 34.

FT DOMAIN 3883 3924
 FT DOMAIN 3928 3966
 FT DOMAIN 3968 4008
 FT DOMAIN 4009 4050
 FT DOMAIN 4156 4197
 FT DOMAIN 4199 4241
 FT DOMAIN 4244 4284
 FT DOMAIN 4332 4370
 FT DOMAIN 4379 4413
 FT SITE 4454 4460
 FT SITE 4457 4463
 FT SITE 4606 4609
 FT SITE 4619 4625
 FT SITE 4624 4630
 FT SITE 1743 1745
 FT SITE 4522 4527
 FT SITE 4601 4606
 FT DISULFID 28 40
 FT DISULFID 35 53
 FT DISULFID 47 62
 FT DISULFID 67 80
 FT DISULFID 74 93
 FT DISULFID 87 103
 FT DISULFID 108 120
 FT DISULFID 115 133
 FT DISULFID 127 142
 FT DISULFID 147 157

LDL-RECEPTOR CLASS A 35.
 LDL-RECEPTOR CLASS A 36.
 EGF-LIKE 14.
 EGF-LIKE 15. CALCIUM-BINDING (POTENTIAL).
 LDL-RECEPTOR CLASS B 35.
 LDL-RECEPTOR CLASS B 36.
 LDL-RECEPTOR CLASS B 37.
 EGF-LIKE 16.
 EGF-LIKE 17.
 SH3-BINDING (POTENTIAL).
 SH3-BINDING (POTENTIAL).
 SH2-BINDING (POTENTIAL).
 SH3-BINDING (POTENTIAL).
 SH3-BINDING (POTENTIAL).
 CELL ATTACHMENT SITE (POTENTIAL).
 ENDOCYTOSIS SIGNAL (POTENTIAL).
 ENDOCYTOSIS SIGNAL (POTENTIAL).
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.

Query Match 17.6%; Score 57.5; DB 1; Length 4660;
 Best Local Similarity 34.1%; Pred. No. 1.7e+02;
 Matches 14; Conservative 4; Mismatches 18; Indels 5; Gaps 2;

QY 14 GSCPIILIRCAMLPNRCLKDT-DCPGIKKCGSCGMAC 53
 Db 1148 GTCQPTQFRC----PDHRCISPLYVCDKDCADGDEAGC 1184

RESULT 59

M3KA HUMAN
 ID M3KA HUMAN STANDARD; PRT; 954 AA.
 AC Q02779; Q12761; Q14871;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)
 DE (Mixed lineage kinase 2) (Protein kinase MST).
 GN MAP3K10 OR MLK2 OR MST.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96128179; PubMed=8536694;
 RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
 RA Sutherland G.R., Simpson R.J.;
 RT "Complete nucleotide sequence, expression, and chromosomal
 RT localisation of human mixed-lineage kinase 2.";
 RL Eur. J. Biochem. 234:492-500(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95249256; PubMed=7731697; Terada M.;
 RA Katoh M., Hirai M., Sugimura T.;
 RT "Cloning and characterization of MST, a novel (putative)
 RT serine/threonine kinase with SH3 domain.";
 RL Oncogene 10:1447-1451(1995).
 RN [3]
 RP SEQUENCE OF 244-480 FROM N.A.
 RC TISSUE=Colon epithelium;
 RX MEDLINE=93238756; PubMed=8477742;
 RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;

RT "Identification of a new family of human epithelial protein kinases
RL containing two leucine/leucine-zipper domains.";
CC Eur. J. Biochem. 213:701-710(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X90846; CAA62351.1; -;
CC EMBL; Z48615; CAA88531.1; -;
CC PIR; S32468; S32468.
CC HSSP; P11362; 1FGK.
CC Genew; HGNC:6849; MAP3K10.
CC MIM; 600137; -;
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00018; SH3; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00452; SH3DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_pkinase; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00326; SH3; 1.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
CC PROSITE; PS50002; SH3; 1.
CC Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; SH3 domain.
FT DOMAIN 2 5 POLY-GLU.
FT DOMAIN 16 81 SH3.
FT DOMAIN 98 360 PROTEIN KINASE.
FT NP_BIND 104 112 ATP (BY SIMILARITY).
FT BINDING 125 125 ATP (BY SIMILARITY).
FT ACT_SITE 222 222 BY SIMILARITY.
FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).
FT CONFLICT 462 464 SRL -> AV (IN REF. 2).
FT CONFLICT 465 480 LKLRGGSHSLPSGF -> AQAAGRQPHQPALWL (IN REF. 3).
FT CONFLICT 471 471 G -> S (IN REF. 2).
FT CONFLICT 807 807 G -> R (IN REF. 2).
FT CONFLICT 818 818 V -> A (IN REF. 2).
FT SEQUENCE 954 AA; 103623 MW; 538F4AA55980A8A CRC64;
Query Match 17.4%; Score 57; DB 1; Length 954;
Best Local Similarity 34.9%; Pred. No. 52;
Matches 15; Conservative 3; Mismatches 13; Indels 12; Gaps 1;
QY 3 EPVKGPVSTKPGSCP-----IILRCAMLNPNRCL 33
DB 130 DPEKPAVTAEQVCQEARLFGALQHPNIIALRGACLNPHLCL 172
RESULT 60
ID IELL_MOMCH STANDARD; PRT; 28 AA.
AC P10286;
DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Elastase inhibitor I (MCEI-I).
OS Momordica charantia (Bitter gourd) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=89291812; PubMed=2738047;
RA Hara S., Makino J., Ikenaka T.;
RT "Amino acid sequences and disulfide bridges of serine proteinase
inhibitors from bitter gourd (Momordica charantia LINN.) seeds.";
RL J. Biochem. 105:88-92(1989).
CC -1- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
CC PIR; JX0059; JX0059.
DR HSSP; P01074; 1CTI.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHETR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; PTI; 1.
DR PROSITE; PS00286; SQUASH INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT_SITE 5 6 REACTIVE BOND.
FT DISULFID 3 20 BY SIMILARITY.
FT DISULFID 10 22 BY SIMILARITY.
FT DISULFID 16 27 BY SIMILARITY.
SQ SEQUENCE 28 AA; 3169 MW; 961C63D22C74FEC8 CRC64;
Query Match 17.3%; Score 56.5; DB 1; Length 28;
Best Local Similarity 31.4%; Pred. No. 3.2;
Matches 11; Conservative 5; Mismatches 10; Indels 9; Gaps 1;
QY 16 CPILIRCAMLNPNRCLKDTCPGKICGCGSCG 50
DB 3 CPLIWCEK-----RSDCLAQICVDGHCG 28
RESULT 61
MTIB HUMAN STANDARD; PRT; 61 AA.
ID MTIB_HUMAN
AC P07438;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-IB (MT-IB).
GN MTIB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87064506; PubMed=3785191;
RA Heguy A., West A., Richards R.I., Karin M.;
RT "Structure and tissue-specific expression of the human
metallothionein IB gene.";
RL Mol. Cell. Biol. 6:2149-2157(1986).
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
CC GLUCOCORTICOID.
CC -1- SUBUNIT: MONOMER.
CC -1- DOMAIN: CLASS 1 METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M13485; AAA36331.1; -
 CC EMBL; M13484; AAA36331.1; JOINED.
 CC PIR; A25244; A25244.
 CC HSP; P02795; 2WHU.
 CC Genew; HGNC:7394; MTLB.
 CC MIN; 156349; -
 CC InterPro; IPR003019; Metallthion.
 CC InterPro; IPR000006; Metallthion_vert.
 CC Pfam; PF00131; metalthio; 1
 CC PRINTS; PR00860; MVTEBRATE.
 CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
 CC Metal-binding; Metal-thiolate cluster; Zinc; Copper; Cadmium;
 CC Acetylation.
 CC MOD RES 1 1 ACETYLATION (BY SIMILARITY).
 CC DOMAIN 1 29 BETA.
 CC DOMAIN 30 61 ALPHA.
 CC METAL 5 5 CLUSTER B.
 CC METAL 7 7 CLUSTER B.
 CC METAL 13 13 CLUSTER B.
 CC METAL 15 15 CLUSTER B.
 CC METAL 19 19 CLUSTER B.
 CC METAL 21 21 CLUSTER B.
 CC METAL 24 24 CLUSTER B.
 CC METAL 26 26 CLUSTER B.
 CC METAL 29 29 CLUSTER B.
 CC METAL 33 33 CLUSTER A.
 CC METAL 34 34 CLUSTER A.
 CC METAL 36 36 CLUSTER A.
 CC METAL 37 37 CLUSTER A.
 CC METAL 41 41 CLUSTER A.
 CC METAL 44 44 CLUSTER A.
 CC METAL 48 48 CLUSTER A.
 CC METAL 50 50 CLUSTER A.
 CC METAL 57 57 CLUSTER A.
 CC METAL 59 59 CLUSTER A.
 CC METAL 60 60 CLUSTER A.
 CC SEQUENCE 61 AA; 6115 MW; 8DC566E5462810D3 CRC64;

Query Match 17.3%; Score 56.5; DB 1; Length 61;
 Best Local Similarity 34.1%; Pred. No. 6.2;
 Matches 15; Conservative 2; Mismatches 16; Indels 11; Gaps 3;
 QY 11 TKPGSCPILIRCAMLPPNRC-LKDTDCPGIKKCCGCGMAC 53
 Db 8 TTGSC-----ACA-----GSCCKEKCCTISCKKCCSCCPVC 41

RESULT 62
 RL25_CHLTR STANDARD; PRT; 185 AA.
 AC O84805;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable 50S ribosomal protein L25.
 GN RPL5 OR CT799.
 OS Chlamydia trachomatis
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 CC NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/JW-3/Ck;
 RX MEDLINE=99000809; PubMed=97841136;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,

RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis";
 RL Science 282:754-759(1998).
 CC -!- FUNCTION: BINDS TO THE 5S RNA (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE L25P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE001352; AAC68394.1; -
 CC PHCI-2DPAGE; O84805; -
 CC InterPro; IPR001021; Ribosomal L25.
 CC Pfam; PF01386; Ribosomal L25p; 1.
 CC TIGRFAMs; TIGR00731; ctc-TL5; 1.
 CC Ribosomal protein; rRNA-binding; Complete proteome.
 CC SEQUENCE 185 AA; 20439 MW; 0E88BBE105B71547 CRC64;

Query Match 17.3%; Score 56.5; DB 1; Length 185;
 Best Local Similarity 31.8%; Pred. No. 15;
 Matches 14; Conservative 3; Mismatches 18; Indels 9; Gaps 1;
 QY 22 RCAMLNPPNRC-LKDTDCPGIK-----KCCEGSCGCMACFVP 56
 Db 100 RDVRLNIPRCINTVDCVGLGSLRQVIRCVWCKPKDVP 143

RESULT 63
 YAW5_SCHPO STANDARD; PRT; 326 AA.
 AC Q10180;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C3F10.05c in chromosome 1.
 GN SPAC3F10.05C.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Calibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

```
RA Cerrutti L., Lowe T., McCombie W.R., Pauleen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69369; CAA93303.1; -
KW Hypothetical protein.
SQ SEQUENCE 326 AA; 37867 MW; 7D88277601A571BC CRC64;
Query Match 17.3%; Score 56.5; DB 1; Length 326;
Best Local Similarity 34.2%; Pred. No. 24;
Matches 13; Conservative 7; Mismatches 17; Indels 1; Gaps 1;
QY 4 PVKGVSTKPGSCPIILIRCAMLPNPNRLKDT-DCPG 40
DB 164 PTRGVVLEWDINPLLSERCKLLLOSELNCQSDYDCPG 201
RESULT 64
CX2X_CONBE STANDARD; PRT; 70 AA.
AC Q9U3Z3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Conotoxin BcTX precursor.
OS Conus betulinus (Beech cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89764;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Venom duct;
RA Fan C.-X.;
RT "A conotoxin with novel cysteine framework from Conus betulinus of
RT south China sea.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF208661; AAF23167.1; -
KW Neurotoxin; Toxin; Signal; Amidation;
KW Cleavage on pair of basic residues.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 57 CONOTOXIN BETX.
FT PROPEP 61 70
FT MOD_RES 57 57
FT SEQUENCE 70 AA; 7900 MW; F6575A2E830AD903 CRC64;
Query Match 17.1%; Score 56; DB 1; Length 70;
Best Local Similarity 24.5%; Pred. No. 7.8;
Matches 13; Conservative 10; Mismatches 22; Indels 8; Gaps 2;
QY 11 TKPGSCPIILIRCAMLPNPNRL-----KDTDCPGIKKCCGSCGCMACFPV 56
DB 6 TSGVCLLLVIVFLNVLVPTISACRAGTYCENDSQ-CLNECCWGGCGHPCRP 57
```

RESULT 65

```
TR18_HUMAN STANDARD; PRT; 241 AA.
ID TR18_HUMAN
AC Q9V5U5; O95851; Q9NYU9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 18 precursor
DE (Glucocorticoid-induced TNFR-related protein) (Activation-inducible
DE TNFR family receptor).
GN TNFRSF18 OR GTR OR AITR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Bone marrow;
RX MEDLINE=99175482; PubMed=10074428;
RA Gurney A.L., Marsters S.A., Huang R.M., Pitti R.M., Mark D.T.,
RA Baldwin D.T., Gray A.M., Dowd A.D., Brush A.D., Heldens A.D.,
RA Schow A.D., Goddard A.D., Wood W.I., Baker K.P., Godowski P.J.,
RA Ashkenazi A.;
RT "Identification of a new member of the tumor necrosis factor family
RT and its receptor, a human ortholog of mouse GTR.";
RL Curr. Biol. 9:215-218(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=T-cell;
RX MEDLINE=99156876; PubMed=10037686;
RA Kwon B., Yu K.-Y., Ni J., Yu G.-L., Jang I.-K., Kim Y.-J., Xing L.,
RA Liu D., Wang S.-X., Kwon B.-S.;
RT "Identification of a novel activation-inducible protein of the tumor
RT necrosis factor receptor superfamily and its ligand.";
RL J. Biol. Chem. 274:6056-6061(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Thymus;
RA Nocentini G., Bartoli A., Ronchetti S., Mastrodicasa E., Riccardi C.;
RT "Identification of a soluble human GTR splicing (hGTR-D).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR TNFSF18. SEEMS TO BE INVOLVED IN
CC INTERACTIONS BETWEEN ACTIVATED T LYMPHOCYTES AND ENDOTHELIAL CELLS
CC AND IN THE REGULATION OF T CELL RECEPTOR-MEDIATED CELL DEATH.
CC MEDIATED NF-KAPPA-B ACTIVATION VIA THE TRAF2/NIK PATHWAY.
CC -!- SUBUNIT: BINDS TO TRAF1 (TNF RECEPTOR-ASSOCIATED FACTOR 1), TRAF2,
CC AND TRAF3 BUT NOT TRAF5 AND TRAF6.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);
CC Secreted (isoform 2).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2/GTR-D; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN LYMPH NODE, PERIPHERAL BLOOD
CC LEUKOCYTES AND WEAKLY IN SPLEEN.
CC -!- INDUCTION: UPREGULATED IN PERIPHERAL MONONUCLEAR CELLS AFTER
CC ANTIGEN STIMULATION/LYMPHOCYTE ACTIVATION.
CC -!- SIMILARITY: CONTAINS 3 TNFR-CYS REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF125304; AAD22635.1; -
DR EMBL; AF117297; AAD18694.1; -
DR EMBL; AF241229; AAF63506.1; -
DR Genew; HGNC:11914; TNFRSF18.
DR MIM; 603905; -
DR InterPro; IPR001368; TNFR_c6.
```


-1- FUNCTION:- INHIBITS GASTROINTESTINAL MOTILITY AND GASTRIC ACID SECRETION. COULD FUNCTION AS A STRUCTURAL COMPONENT OF GASTRIC MUCUS, POSSIBLY BY STABILIZING GLYCOPROTEINS IN THE MUCUS GEL THROUGH INTERACTIONS WITH CARBOHYDRATE SIDE CHAINS.

-1- SUBCELLULAR LOCATION: Secreted.

-1- TISSUE SPECIFICITY: EXPRESSED IN THE DIGESTIVE TRACT, WHERE IT WAS FOUND PREDOMINANTLY IN THE STOMACH WITH HIGHEST EXPRESSION IN THE ANTRUM. IT IS SECRETED PREDOMINANTLY FROM ANTRAL MUCOUS CELLS INTO THE LUMEN OF THE GASTROINTESTINAL DOMAIN.

-1- SIMILARITY: CONTAINS 2 P-TYPE (TREFOIL) DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).

EMBL; M97255; AAA19025.1; ..
HSSP; P01359; LPCP.
InterPro; IPR000519; P_trefoil.
Pfam; PF00088; trefoil; 2.
PRINTS; PR00680; PTREFOIL.
SMART; SM00018; P; 2.
DR PROSITE; PS00025; P_TREFOIL; 2.
KW Repeat; Growth factor; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 129 TREFOIL FACTOR 2.
FT DOMAIN 80 121 P-TYPE 1.
FT DOMAIN 80 121 P-TYPE 2.
FT DISULFID 29 127 BY SIMILARITY.
FT DISULFID 31 58 BY SIMILARITY.
FT DISULFID 42 57 BY SIMILARITY.
FT DISULFID 52 69 BY SIMILARITY.
FT DISULFID 81 107 BY SIMILARITY.
FT DISULFID 91 106 BY SIMILARITY.
FT DISULFID 101 118 BY SIMILARITY.
SQ SEQUENCE 129 AA; 14077 MW; 00SD0848F64F92D2 CRC64;

Query Match 17.0%; Score 55.5; DB 1; Length 129;
Best Local Similarity 34.5%; Pred.No. 15;
Matches 19; Conservative 4; Mismatches 15; Indels 17; Gaps 5;

QY 12 KPGSCPIILIRCAMLNPPNRCLKLDTCGKIK-----CCEGS--CGWA-CFVP 56
||| : | | | | | ||:
DB 25 KPSPC----RCSRWTPSNR--KNCGFPGITSDQCFLNLCCTFSVAGVPWC FHP 72

RESULT 72

ID TMPEP HUMAN STANDARD; PRT; 287 AA.

AC Q96SW9; O9UJD3; Q96B72;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Transmembrane prostate androgen-induced protein (Solid tumor-associated i1 protein).
DE DE TMPEAI OR PMEPAl OR STAgI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_taxid=9606;
[1]
RN RN
RP RP
RS RS
SEQINFO FROM N.A. (ISOFORM 2).
MEDLINE=20334621; PubMed=10873380;
Xu L.L., Shanmugan N., Segawa T., Seesterhenn I.A., McLeod D.G., Maul J.W., Srivastava S.;
RA "A novel androgen-regulated gene, PMEPAl, located on chromosome 20q13 exhibits high level expression in prostate.";
RL Genomics 66:257-263(2000).
[2]
RN RN
RP RP
RS RS
SEQINFO FROM N.A. (ISOFORM 1).

RA MEDLINE=21453682; PubMed=11568975;
RA Rae F.K., Hooper J.D., Nicol D.L., Clements J.A.;
RT "Characterization of a novel gene, STAG1/PMEP1, upregulated in renal
RT cell carcinoma and other solid tumors.";
RL Mol. Carcinog. 32:44-53(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RX Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.D., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [4]
RP SEQUENCE OF 8-287 FROM N.A. (ISOFORM 1).
RX TISSUE=Kidney;
RC Strausberg R.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Interacts with the WW domains of NEDD4 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: Highest expression in prostate. Also expressed
CC in ovary.
CC -!- INDUCTION: By androgen.
CC -!- SIMILARITY: BELONGS TO THE TMEPAI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR ENBL; AF224278; AAF86322.1; -
DR ENBL; AF305616; AAL16781.1; -
DR ENBL; AF305426; AAL09357.1; -
DR ENBL; AL035541; CAB55862.1; -
DR ENBL; BC015918; AAH15918.1; ALT_INIT.
DR GenBank; HGNC:14107; TMEPAI.
DR MTM; 606564; -
RW Transmembrane; Alternative splicing.
KW DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 63 POTENTIAL.
FT DOMAIN 64 287 CYTOPLASMIC (POTENTIAL).
FT SITE 158 161 WW-BINDING (POTENTIAL).
FT SITE 229 232 WW-BINDING (POTENTIAL).
FT VARSPPLIC 1 37 MRLMGVNSTAAAGAPNVSCCTCNKRSLFQSMETIT ->
FT MA (IN ISOFORM 2).
FT SEQUENCE 287 AA; 31609 MW; 6103473561AE08DA CRC64;

Query Match 17.0%; Score 55.5; DB 1; Length 287;
Best Local Similarity 29.7%; Pred. No. 28;
Matches 19; Conservative 8; Mismatches 20; Indels 17; Gaps 3;
QY 4 PVKGP-VSTKPGSCPIILIRCAMLNPNRCLKQTD-----CP-----GIKKCCCE 46
DB 160 PYQGPTQLRLDPEOQLNRESVRAPNRTIFDSLDMSARLQGPCPPSSNGISATCY 219
QY 47 GSGC 50
DB 220 GSGC 223
RESULT 73
U713 HUMAN STANDARD; PRT; 500 AA.
ID Q9BYM8; O95623; Q9BYM9;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin conjugating enzyme 7 interacting protein 3 (Hepatitis B
DE virus X-associated protein 4) (HBV associated factor 4).
GN UBCEV1P3 OR C200RF18 OR XAP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RA Seto E.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., King A., Knights A., Laird G.K., Lawlor S.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Lloyd D.M., Lovell J.D.,
RA Levaslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., McMurray A.A.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., Nickerson T.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.D., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RX TISSUE=Placenta;
RC Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
CC OF THE E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2
CC UBIQUITIN-CONJUGATING ENZYMES, SUCH AS UBE2L3/UBC4, AND THEN
CC TRANSFERS IT TO SUBSTRATES.
CC -!- SUBUNIT: INTERACTS WITH BETA-1-TYPE (PRKCB1) AND ZETA-TYPE PROTEIN
CC KINASE C (PRKCZ) AND WITH UBE2L3. FORMS HOMODIMERS IN VITRO (BY


```
CC SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1/RCK1 (SHOWN HERE), 2/RBCK2, 3
CC AND 4; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; U67322; AAD00162.1; -.
DR EMBL; AL121747; CAC17516.1; -.
DR EMBL; AL121747; CAC28311.1; -.
DR EMBL; AL121747; CAC28312.1; -.
DR EMBL; BC000983; AAH00983.3; -.
DR Genew; HGNC:15864; C20orf18.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR000315; ZnF_Box.
DR InterPro; IPR001876; ZnF_RanGDP.
DR InterPro; IPR001841; ZnF_Ring.
DR Pfam; PF00097; zf-C3HC4_1.
DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00547; ZnF_RBZ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR PROSITE; PS50119; ZF_BOX; FALSE NEG.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS01199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR UniProt conjugation pathway; Zinc-finger; Coiled coil; Phosphorylation;
KW Alternative splicing.
FT DOMAIN 45 109 UBIQUITIN-LIKE.
FT FT ZN FING 183 212 RANBP2-TYPE.
FT FT DOMAIN 223 251 COILED COIL (POTENTIAL).
FT FT ZN FING 272 317 RING-TYPE.
FT FT ZN FING 366 401 B BOX-TYPE.
FT FT ZN FING 437 463 B BOX-LIKE-TYPE.
FT FT VARSPLIC 1 45 MALSLTRAVAGGDQVAMKCAIWLAEQRVPLSVOLKPVEVS
FT TQDI -> MGTATPDGREQE (IN ISOFORM 3 AND
FT ISOFORM 4).
FT VARSPLIC 243 262 RKQQQEGNYLVQLDQRS -> GVPAGHHPPQGGLL
FT PLH (IN ISOFORM 2 AND ISOFORM 4).
FT VARSPLIC 263 500 MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT SEQUENCE 500 AA; 56381 MW; 7492E03EB2E9CE4F CRC64;
Query Match 17.0%; Score 55.5; DB 1; Length 500;
Best Local Similarity 30.0%; Pred. No. 45;
Matches 15; Conservative 3; Mismatches 15; Indels 17; Gaps 3;
QY 2 QBPVKGVSTKP-----GSCPILLIRCAMLNPNRCLKDTCDCPGIKKC 45
|||.|.|||:|||:|||||:|||||:|:|:|
Db 169 QBPGRQPDAPPEPPVGWQCP----GCTFINKPTR-----PGCEMCC 207
|||||.|||:|||||:|||||:|:|:|
RESULT 74
VG47 HSV11
ID VG47 HSV11 STANDARD; PRT; 518 AA.
AC Q00139;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 56 kDa serine protease (EC 3.4.21.-).
GN 47.
OC Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
```

```

RL Hoppe-Seyler's Z. Physiol. Chem. 364:93-95(1983).
RN [3]
RP PRELIMINARY SEQUENCE (CMT-III).
RC TISSUE=Seed;
RX MEDLINE=82005824; PubMed=7275008;
RA Nowak K., Slominska A., Polanowski A., Wiczorek M., Wilusz T.;
RT "Trypsin inhibitor III from squash seeds (Cucurbita maxima), its
RT reactive site and proposed amino acid sequence.";
RL Hoppe-Seyler's Z. Physiol. Chem. 362:1017-1019(1981).
RN [4]
RP STRUCTURE BY NMR OF CMT-III.
RX MEDLINE=92118901; PubMed=1731946;
RA Krishnamoorthi R., Gong Y., Lin C.-L.S., Vandervelde D.;
RT "Two-dimensional NMR studies of squash family inhibitors. Sequence-
RT specific proton assignments and secondary structure of reactive-site
RT hydrolyzed Cucurbita maxima trypsin inhibitor III.";
RL Biochemistry 31:898-904(1992).
CC -1- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR PIR; A01313; TIPU3.
DR PIR; S07156; S07156.
DR HSP; P01074; ICTI.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHETR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; FTI; 1.
DR PROSITE; PS00286; SQUASH INHIBITOR; 1.
KW Serine protease inhibitor.
FT CHAIN 1 32 CMTI-IV.
FT CHAIN 4 32 CMTI-III.
FT ACT SITE 8 9 REACTIVE BOND.
FT DISULFID 6 23 BY SIMILARITY.
FT DISULFID 13 25 BY SIMILARITY.
FT DISULFID 19 31 BY SIMILARITY.
SQ SEQUENCE 32 AA; 3669 MW; 0F591120B0137512 CRC64;

Query Match 16.8%; Score 55; DB 1; Length 32;
Best Local Similarity 38.9%; Pred. No. 5.3;
Matches 14; Conservative 3; Mismatches 9; Indels 10; Gaps 2;

QY 16 CPILIRCAMLNPPNRCCLKDTPCGIKKCE-GSCG 50
Db 6 CPRILMKCK-----KSDCLAEVCVLEHGCG 32

```

Search completed: February 11, 2003, 08:37:03
 Job time : 16 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 19:03:02 ; Search time 1295 Seconds
(without alignments)
1280.973 Million cell updates/sec

Title: 09-833799-13b

Perfect score: 327

Sequence: 1 aqepvkgpvtkgscpiil.....cpgikkcccgscgmactvpq 57

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09833799/runat.12022003.120042.23635/app.query.fasta_1.199
-DB=GenEmbl -QMT=fastcap -SUPFIX=rge -MINMATCH=0.1 -LOOPFCU=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-LOCALIGN=200 -THR_SCORE=500 -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799 @CGN 1.1 3745 @runat.12022003.120042.23635 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sc.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	327	100.0	171	6	A31088 DNA fragmen
2	327	100.0	206	6	A31086 DNA fragmen
C 3	327	100.0	206	6	A31087 DNA fragmen
4	327	100.0	347	6	A31089 DNA fragmen
C 5	327	100.0	434	6	AX041087 Sequence
6	327	100.0	478	9	HSANTLEUP
C 7	327	100.0	480	6	AX397178 Sequence
8	327	100.0	504	6	AX31090 DNA fragmen
9	327	100.0	571	6	I74316 Sequence 14
10	327	100.0	582	9	BC010952 Homo sapi
11	327	100.0	1878	9	HUMELAFIN
12	327	100.0	2309	6	AX333644 Sequence
13	327	100.0	2309	6	AX334316 Sequence
14	327	100.0	2309	9	HUMPRELAS
15	327	100.0	2309	9	S58717 pre-elafins
16	327	100.0	99747	9	HSJ172H20
17	324	99.1	177	6	I96047 Sequence 3
18	324	99.1	8598	6	E16036 cDNA encodi
19	323	98.8	177	6	I96048 Sequence 5
20	323	98.8	177	6	I96049 Sequence 7
21	322	98.5	321	6	I74313 Sequence 11
22	250	76.5	573	4	BO223216 Bos tauru
23	221.5	67.7	3693	4	PIGWAPA
24	217.5	66.5	1128	4	AB003281 Phacochoe
25	216.5	66.2	464	4	PIGWAPD
26	216.5	66.2	1034	4	D83668 Sus scrofa
27	201	61.5	1169	4	AB003283 Pecari ta
28	182	55.7	270	9	MAM223215 Macaca mu
29	181.5	55.5	1325	4	AB003282 Phacochoe
30	179.5	54.9	483	4	PIGSPA12S2
31	179.5	54.9	737	6	E06782 cDNA contai
32	179.5	54.9	789	4	DMY245
33	179.5	54.9	3782	4	PIGWAPB
34	177.5	54.3	578	4	PIGWAPB
35	171	52.3	490	10	AF178426 Rattus no
36	170	52.0	495	4	AB011010 Bos tauru
37	166	50.8	734	4	AB003284 Hippopota
38	165	50.5	600	4	PIGALP
39	161	49.2	671	10	MMU88093 Mus musculu
40	161	49.2	682	10	MMU94341 Mus musculu
41	161	49.2	894	10	BC028509 Mus muscu
42	161	49.2	1123	10	MMU73004 Mus muscu
43	160.5	49.1	667	10	AF151982 Rattus no
44	160	48.9	3670	4	PIGWAPC
45	158	48.3	464	10	AB042257 Cavia por
46	156	47.7	194143	2	AC112730 Rattus no
47	148	45.3	2435	10	AF002719 Mus muscu
C 48	148	45.3	190669	10	AL590429 Mouse DNA
49	144	44.0	4090	10	AF205374 Mus muscu
50	143	43.7	194	6	AR067991 Sequence

RESULT 3
A31087/C

```

RESULT 1
A31088
LOCUS       A31088                171 bp    DNA
DEFINITION  DNA fragment from patent EP0402068.
ACCESSION   A31088
VERSION     A31088.1  GI:1249301
KEYWORDS    .
SOURCE      .
ORGANISM    .
REFERENCE   1  (bases 1 to 171)
AUTHORS     Christophers E., Schroder J.M., Pioli D., Wiedow O. and Edge M.D.
TITLE       Polypeptides and polypeptide analogues with inhibitory activity
            against human elastase
JOURNAL     Patent: EP 0402068-A 29 12-DEC-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
            Location/Qualifiers
                1..171
                /organism="synthetic construct"

```

LOCUS A31087 206 bp DNA linear PAT 21-AUG-1995
 DEFINITION DNA fragment from patent EP0402068.
 ACCESSION A31087
 VERSION A31087.1 GI:1247248
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM artificial sequences.
 REFERENCE 1 (bases 1 to 206)
 AUTHORS Christophers, E., Schroder, J.M., Pioli, D., Wiedow, O. and Edge, M.D.
 TITLE Polypeptides and polypeptide analogues with inhibitory activity against human elastase
 JOURNAL Patent: EP 0402068-A 28 12-DEC-1990;
 IMPERIAL CHEMICAL INDUSTRIES PLC
 FEATURES
 source
 1..206
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 64 a 45 c 47 g 50 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8.82e-29 Length: 206
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x A31087 (1-206)
 QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 181 GCTCAAGAACCCAGTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCTATTATCTTG 122
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 121 ATTCGTGGCGCATGTTGAATCCCTTAACCGTGTGTTTGAAGGACACTGATTGCCAGT 62
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 61 ATCAAAAGTGTGTGAAGTCTCTCGGTATGGCTGTGTTGTTCACAA 11

RESULT 4
 A31089
 LOCUS A31089 347 bp DNA linear PAT 21-AUG-1995
 DEFINITION DNA fragment from patent EP0402068.
 ACCESSION A31089
 VERSION A31089.1 GI:1249302
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM artificial sequences.
 REFERENCE 1 (bases 1 to 347)
 AUTHORS Christophers, E., Schroder, J.M., Pioli, D., Wiedow, O. and Edge, M.D.
 TITLE Polypeptides and polypeptide analogues with inhibitory activity against human elastase
 JOURNAL Patent: EP 0402068-A 30 12-DEC-1990;
 IMPERIAL CHEMICAL INDUSTRIES PLC
 FEATURES
 source
 1..347
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 65 a 118 c 78 g 84 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.52e-28 Length: 347
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x A31089 (1-347)
 QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 1 GCGCAAGAGCCAGTCAAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATTAATCTTG 60
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 61 ATCCGGTGGCGCATGTTGAATCCCTTAACCGTGTGTTTGAAGGACACTGATTGCCAGGA 120
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 121 ATAAAGAARTGCTGTGAAGGCTCTTGGGGATGGCGCTGTTTCGTTCGCCAG 171

RESULT 5
 AX041087/c
 LOCUS AX041087 434 bp DNA linear PAT 23-NOV-2000
 DEFINITION Sequence 5 from Patent WO0065053.
 ACCESSION AX041087
 VERSION AX041087.1 GI:11340657
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 434)
 AUTHORS Wang, T. and Dillon, D.C.
 TITLE Compositions and methods for therapy and diagnosis of head/neck and lung squamous cell carcinoma
 JOURNAL Patent: WO 0065053-A 5 02-NOV-2000;
 CORIXA CORPORATION (US)
 FEATURES
 source
 1..434
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 103 a 103 c 133 g 94 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.93e-28 Length: 434
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x AX041087 (1-434)
 QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 318 GCGCAAGAGCCAGTCAAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATTAATCTTG 259
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 258 ATCCGGTGGCGCATGTTGAATCCCTTAACCGTGTGTTTGAAGGACACTGATTGCCAGGA 199
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 198 ATCAAGAARTGCTGTGAAGGCTCTTGGGGATGGCGCTGTTTCGTTCGCCAG 148

RESULT 6
 HSANTLEUP
 LOCUS HSANTLEUP 478 bp mRNA linear PRI 08-JAN-1998
 DEFINITION H.sapiens encoding skin-derived antileukoproteinnase.
 ACCESSION Z18538
 VERSION Z18538.1 GI:28711
 KEYWORDS antileukoproteinnase; elafin; elastase inhibitor; proteinase inhibitor; psoriasis; SKALP; transglutaminase substrate.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 478)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 478)
 AUTHORS Molhuizen,H.O., Alkemade,H.A., Zeeuwen,P.L., de Jongh,G.J.,
 Wieringa,B. and Schalkwijk,J.
 TITLE SKALP/elafin: an elastase inhibitor from cultured human
 keratinocytes. Purification, cDNA sequence, and evidence for
 transglutaminase cross-linking
 J. Biol. Chem. 268 (16), 12028-12032 (1993)
 JOURNAL 93280175
 MEDLINE
 REFERENCE 2 (bases 1 to 478)
 AUTHORS Molhuizen,H.O.F.
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-1992) Molhuizen H.O.F., Academic Hospital,
 Dermatology, Nijmegen, The Netherlands, 6500 HB
 FEATURES
 source
 1. .478
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PGESKA"
 /cell_type="keratinocyte"
 /clone_lib="HL1110b Clonetechn"
 1. .478
 /partial
 /citation=[1]
 3. .356
 /function="protease inhibitor"
 /citation=[1]
 /codon_start=1
 /evidence=experimental
 /product="skin-derived antileukoproteinase"
 /protein_id="CAA7923.1"
 /db_xref="GI:28712"
 /translation="MRASSFLIVVFLIAGTLVLEAAVTGVKGDVTKGRVPFNGQ
 DPVKGVSVKGDVKVAQBPVKGPVSTKPGSCPILIRCAMLNPENRCLKDDCPGIK
 KCCEGSGMACFPVQ"
 3. .68
 /partial
 /product="skin-derived antileukoproteinase"
 /function="protease inhibitor"
 102 a 142 c 119 g 115 t
 BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 2,13e-28 Length: 478
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

09-833799-13B (1-57) x HSANTLEUP (1-478)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 |||||
 Db 183 GCGCAAGGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTGCCCATATCTTG 242
 |||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 243 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCCCAGGA 302
 |||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 303 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGCGCTGTTTCGTTCCCCAG 353
 |||||

RESULT 7
 AX397178/c
 LOCUS AX397178 480 bp DNA linear PAT 18-MAY-2002
 DEFINITION Sequence 1393 from Patent WO0212328.
 ACCESSION AX397178
 VERSION AX397178.1 GI:21067925
 KEYWORDS

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS King,G.E., Meagher,M.J., Xu,J. and Secrist,H.
 TITLE Compositions and methods for the therapy and diagnosis of colon
 cancer
 JOURNAL Patent: WO 0212328-A 1393 14-FEB-2002;
 CORIXA CORPORATION (US)
 FEATURES
 source
 1. .480
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 113 a 105 c 139 g 120 t
 BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 2,14e-28 Length: 480
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x AX397178 (1-480)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 |||||
 Db 369 GCGCAAGGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTGCCCATATCTTG 310
 |||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 309 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCCCAGGA 250
 |||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 249 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGCGCTGTTTCGTTCCCCAG 199
 |||||

RESULT 8
 AX31090
 LOCUS AX31090 504 bp DNA linear PAT 21-AUG-1995
 DEFINITION DNA fragment from patent EP0402068.
 ACCESSION AX31090
 VERSION AX31090.1 GI:1249303
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 KEYWORDS artificial sequences.
 REFERENCE 1 (bases 1 to 504)
 AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
 TITLE Polypeptides and polypeptide analogues with inhibitory activity
 against human elastase
 JOURNAL Patent: EP 0402068-A 31 12-DEC-1990;
 IMPERIAL CHEMICAL INDUSTRIES PLC
 FEATURES
 source
 1. .504
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 110 a 152 c 120 g 122 t
 BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 2,25e-28 Length: 504
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x AX31090 (1-504)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20

```

Db 158 GCGAAGAGCCAGTCAAGGTCAGCTCCACTAAAGCTGGCTCTGCCCCCATTTATCTTG 217
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 218 ATCCGTGGCCCATGTTGAATCCCTTAACCGCTGTTTGAAGATACACTGACTGCCAGGA 277
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 278 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTGGTCCCCAG 328

RESULT 9
I74316
LOCUS I74316 571 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 14 from patent US 5688641.
ACCESSION I74316
VERSION I74316.1 GI:3010457
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 571)
AUTHORS Sager R., Zou Z., Lee S. Whan, and Tomasetto C. Laure.
TITLE Cancer diagnosis using nucleic acid hybridization
JOURNAL Patent: US 5688641-A 14 18-NOV-1997;
FEATURES
Location/Qualifiers
1..571
/organism="unknown"
BASE COUNT 138 a 167 c 130 g 136 t
ORIGIN

Alignment Scores:
Pred. No.: 2,576-28 Length: 571
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

09-833799-13B (1-57) x I74316 (1-571)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
Db 199 GCGAAGAGCCAGTCAAGGTCAGCTCCACTAAAGCTGGCTCTGCCCCCATTTATCTTG 258
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 259 ATCCGTGGCCCATGTTGAATCTCTTAACCGCTGTTTGAAGATACACTGACTGCCAGGA 318
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 319 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTGGTCCCCAG 369

RESULT 10
BC010952
LOCUS BC010952 582 bp mRNA linear PRI 25-JUL-2001
DEFINITION Homo sapiens, Similar to protease inhibitor 3, skin-derived
(SKALP), clone MGC:13613 IMAGE:4083155, mRNA, complete cds.
ACCESSION BC010952
VERSION BC010952.1 GI:15012094
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Strausberg, R.
Direct Submission
Submitted (23-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK

```

COMMENT

Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 19 Row: d Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505786.

FEATURES

Location/Qualifiers
1..582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:13613 IMAGE:4083155"
/tissue_type="Brain, glioblastoma"
/clone_lib="NIH MGC_57"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
26..379
/codon_start=1
/product="Similar to protease inhibitor 3, skin-derived
(SKALP)"
/protein_id="AAH10952.1"
/db_xref="GI:15012095"
/translation="MRASSFLIVVFLIAGTLVLEAAVGVKVGQDTLVKGRVPPNGQ
DPVKGVSVKGDQKVKAPQPVKSPVSTKPSGCPILIRCAMLPPNRLKDTDCPGIK
KCEGSCGMACFPQ"

CDS

148 a 166 c 134 g 134 t
ORIGIN

Alignment Scores:

Pred. No.: 2,626-28 Length: 582
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x BC010952 (1-582)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20

Db 206 GCGAAGAGCCAGTCAAGGTCAGCTCCACTAAAGCTGGCTCTGCCCCCATTTATCTTG 265

Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 266 ATCCGTGGCCCATGTTGAATCTCTTAACCGCTGTTTGAAGATACACTGACTGCCAGGA 325

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 326 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTGGTCCCCAG 376

RESULT 11

HUMELAFIN

LOCUS HUMELAFIN 1878 bp DNA linear PRI 14-APR-2000
DEFINITION Human gene for elafin, complete cds.

ACCESSION

D13156

VERSION D13156.1 GI:219614

KEYWORDS elafin; elastase inhibitor.

SOURCE

Homo sapiens (library: lambda EMBL DNA, clone lambda hi-G1.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1878)
Saheki, T., Ito, F., Hagiwara, H., Saito, Y., Kuroki, J., Tachibana, S.
and Hirose, S.


```

repeat_region /note="L1M4b repeat: matches 330. .524 of consensus"
22891. .23199
repeat_region /note="AluX repeat: matches 1. .309 of consensus"
25460. .25629
repeat_region /note="L1R16A repeat: matches 201. .378 of consensus"
27612. .27639
repeat_region /note="14 copies 2 mer ac 100% conserved"
27655. .27986
repeat_region /note="L1M3a repeat: matches 202. .531 of consensus"
28959. .29336
repeat_region /note="L1M1 repeat: matches 469. .859 of consensus"
29337. .29711
repeat_region /note="THE1C repeat: matches 1. .371 of consensus"
29712. .31317
repeat_region /note="THE1C-internal repeat: matches 1. .1580 of consensus"
31318. .31729
repeat_region /note="THE1C repeat: matches 1. .371 of consensus"
31730. .31886
repeat_region /note="L1M1 repeat: matches 859. .987 of consensus"
33209. .34261
repeat_region /note="L1M3d repeat: matches 859. .1833 of consensus"
34232. .34942
repeat_region /note="L1M1 repeat: matches 1560. .2657 of consensus"
34941. .36117
repeat_region /note="L1 repeat: matches 3935. .5142 of consensus"
36118. .36412
repeat_region /note="AluJo repeat: matches 9. .292 of consensus"
36430. .37302
repeat_region /note="L1 repeat: matches 2905. .3780 of consensus"
37299. .37425
repeat_region /note="L1 repeat: matches 4912. .5038 of consensus"
37434. .37469
repeat_region /note="L1P repeat: matches 2850. .2885 of consensus"
38377. .39183
repeat_region /note="L1M4c repeat: matches 1312. .1666 of consensus"
39261. .42088
repeat_region /note="L1P1A5 repeat: matches 671. .3485 of consensus"
42057. .43098
repeat_region /note="L1P1A5 repeat: matches 5104. .6143 of consensus"
43119. .45775
repeat_region /note="L1P1A5 repeat: matches 3467. .6157 of consensus"
45840. .45956
repeat_region /note="L1M1E repeat: matches 5530. .5647 of consensus"
46081. .46114
repeat_region /note="17 copies 2 mer tc 82% conserved"
46492. .46761
repeat_region /note="L1M1E repeat: matches 5521. .5793 of consensus"
46828. .46915
repeat_region /note="L1M1E3A repeat: matches 5300. .5388 of consensus"
46916. .47210
repeat_region /note="AluSq repeat: matches 1. .305 of consensus"
47211. .47769
repeat_region /note="L1M1E3A repeat: matches 5388. .5937 of consensus"
complement(48203. .48496)
repeat_region /note="match: GSS: Em:AQ275745"
48638. .48679
gene /note="21 copies 2 mer ac 97% conserved"
49535. .51104
mRNA /gene="PI3"
join(49535. .49615,50475. .50750,50984. .51104)
/product="dJ172H20.3 (protease inhibitor 3, skin-derived (SKALP))"
match: cDNAs: Em:Z18538 Em:AJ223218 Em:AJ223216
match: ESTs: Em:AA527557 Em:AA583567 Em:AA582866
evidence-not experimental
join(49537. .49615,50475. .50749)

CDs

Alignment Scores:
Pred. No.: Length:
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x HSJ172H20 (1-99747)

QY 1 AlaGlnGluProVallysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 50576 GCGCAAGAGCAGTCAAAAGTCCAGTCTCCACTAAGCTGGCTCTTGGCCCATTAATCTTG 50635

QY 21 IleArgCysAlaMetIleuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 50636 ATCCGCTGCCCATGTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCAGGA 50695

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 50696 ATCAAGAAGTGCTGTGAAGGCTCTTGGCGATGGCTGTTCGTTCCCGCAG 50746

RESULT 17
LOCUS 196047
DEFINITION Sequence 3 from patent US 5734014.
ACCESSION 196047
VERSION 196047.1 GI:3940517
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaji,A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 3 31-MAR-1998;
FEATURES
Source 1..177
BASE COUNT 39 a 49 c 48 g 41 t
ORIGIN

Alignment Scores:
Pred. No.: Length:
Score: 1.65e-28 Matches: 177
Percent Similarity: 324.00% Conservative: 56
Best Local Similarity: 100.00% Mismatches: 1
Query Match: 98.25% Indels: 0
DB: 99.08% Gaps: 0

09-833799-13B (1-57) x 196047 (1-177)

QY 1 AlaGlnGluProVallysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 1 GCACAGGAACCAAGTTAAAGTCCGGTGTGCGACCAACCGGCTCTTGGCCGATTATCTTG 60

QY 21 IleArgCysAlaMetIleuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 61 ATCCGCTGCCCTTTGCTGAACCCCGCGAACCGTGTCTGTAAGACACTGACTGCCGGGT 120

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 121 ATCAAAAATGCTGGAAGTTCTTTCGGGTATGGCATGCTTCGTTCCCGCAG 171

RESULT 18
LOCUS E16036
DEFINITION cDNA encoding improved Pichia elafin.
ACCESSION E16036
VERSION E16036.1 GI:5710719
KEYWORDS JP 199812792-A/1.
SOURCE Pichia pastoris.
ORGANISM Pichia pastoris
REFERENCE Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Saccharomycetales; Saccharomycetaceae; Pichia.
1 (bases 1 to 8598)
Taniyama,M., Yamamoto,T., Okawa,N. and Zushi,M.

```

TITLE ELAFINS-EXPRESSION VECTOR AND PRODUCTION OF ELAFINS BY UTILIZING
 JOURNAL THE SAME
 COMMENT Patent: JP 1998127292-A 1 19-MAY-1998;
 TSUMURA & CO
 OS Pichia pastoris
 PN JP 1998127292-A/1
 PD 19-OCT-1998
 PF 31-OCT-1996 JP 1996304233
 PI TANIYAMA MITSUE, YAMAMOTO TAKASHI, OKAWA NORIYUKI, PI ZUSHI
 MAKOTO
 PC C12N15/09.C07H21/04.C07K14/39.C12N1/19.C12P21/02.(C12N1/19, PC
 C12R1.84),
 PC (C12P21/02.C12R1.84);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No; Location/Qualifiers
 FH Key
 FH
 FT source 1..8598
 FT /organism='Pichia pastoris'
 FT /clone='pPIC9/ELP25L'
 FT 949..1377
 FT /product='improved Pichia elafin' FT
 sig_peptide 949..1203
 FT mat_peptide 1204..1374
 FT misc_feature replace(1276..1278,'atg')
 FT /note='native Pichia elafin'.
 FT Location/Qualifiers
 1..8598
 /organism='Pichia pastoris'
 /db xref='taxon:4922'
 BASE COUNT 2306 a 2025 c 1920 g 2347 t
 ORIGIN

Alignment Scores:
 Pred. No.: 9.69e-27 Length: 8598
 Score: 324.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 99.08% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x E16036 (1-8598)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProGlyLeu 20
 Db 1204 GCTCAAGAACCAAGTTAAGGTCGGTTCGACCAACCGGGCTCTTGCCTGATATCTG 1263

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 1264 ATCCGCTGGCTTTCGTAACCGCGCAACCGCTTCTCTGAAAGACACTGACTGCCGGGT 1323

QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 1324 ATCAAAAATGCTGCAAGGTTCTTTCGGGTATGGCATGCTTCGTTCCCGAG 1374

RESULT 19
 LOCUS 196048 177 bp DNA linear PAT 01-DEC-1998
 DEFINITION Sequence 5 from patent US 5734014.
 ACCESSION 196048
 VERSION 196048.1 GI:3940518
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES
 1 (bases 1 to 177)
 AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kajii,A.
 TITLE Elafin derivative
 JOURNAL Patent: US 5734014-A 5 31-MAR-1998;
 FEATURES Location/Qualifiers

source 1..177
 /organism='unknown'
 BASE COUNT 39 a 50 c 48 g 40 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.14e-28 Length: 177
 Score: 323.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x I96049 (1-177)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProGlyLeu 20
 Db 1 GCACAGAACCAAGTTAAAGTCCGGTTCGACCAACCGGGCTCTTGCCTGATATCTG 60

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 61 ATCCGCTGGCTTTCGTAACCGCGCAACCGCTTCTCTGAAAGACACTGACTGCCGGGT 120

QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 121 ATCAAAAATGCTGCAAGGTTCTTTCGGGTATGGCATGCTTCGTTCCCGAG 171

RESULT 20
 LOCUS 196049 177 bp DNA linear PAT 01-DEC-1998
 DEFINITION Sequence 7 from patent US 5734014.
 ACCESSION 196049
 VERSION 196049.1 GI:3940519
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES
 1 (bases 1 to 177)
 AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kajii,A.
 TITLE Elafin derivative
 JOURNAL Patent: US 5734014-A 7 31-MAR-1998;
 FEATURES Location/Qualifiers

source 1..177
 /organism='unknown'
 BASE COUNT 40 a 49 c 47 g 41 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.14e-28 Length: 177
 Score: 323.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0

09-833799-13B (1-57) x I96049 (1-177)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProGlyLeu 20
 Db 1 GCACAGAACCAAGTTAAAGTCCGGTTCGACCAACCGGGCTCTTGCCTGATATCTG 60

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 61 ATCCGCTGGCTTTCGTAACCGCGCAACCGCTTCTCTGAAAGACACTGACTGCCGGGT 120

QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 121 ATCAAAAATGCTGCAAGGTTCTTTCGGGTATGGCATGCTTCGTTCCCGAG 171

RESULT 21
 LOCUS 174313 321 bp DNA linear PAT 03-APR-1998
 DEFINITION Sequence 11 from patent US 5688641.

```

ACCESSION I74313
VERSION I74313.1 GI:3010454
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 321)
AUTHORS Sager, R., Zou, Z., Lee, S. Whan, and Tomasetto, C. laure.
TITLE Cancer diagnosis using nucleic acid hybridization
JOURNAL Patent: US 568641-A 11 18-NOV-1997;
MEDLINE Location/Qualifiers
FEATURES
1..321
/organism="unknown"
BASE COUNT 80 a 80 c 85 g 76 t
ORIGIN

Alignment Scores:
Pred. No.: 5,2e-28 Length: 321
Score: 322.00 Matches: 56
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.47% Indels: 0
DB: 6 Gaps: 0

09-833799-13B (1-57) x I74313 (1-321)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 144 GCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCGCTGGCTCTGCCCATTCCTTG 203
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 204 ATCCGTGGCCCATGTGAATCCTCTCAACCGCTGCTTGAAGAGATACTGACTGCCACGA 263
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 264 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCTCC 311

RESULT 22
LOCUS BOT223216 573 bp mRNA linear MAM 15-DEC-1998
DEFINITION Bos taurus mRNA for putative bTrappin-2 protein, partial.
ACCESSION AJ223216
VERSION AJ223216.1 GI:2764785
KEYWORDS bTrappin-2 gene; proteinase inhibitor, putative; transglutaminase
SUBSTRATE.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 573)
Zeeuwen, P.L.J.M.
Direct Submission
Submitted (06-JAN-1998) Zeeuwen P.L.J.M., Dermatology, University
Hospital Nijmegen, PO BOX 9101, 6500 HB Nijmegen, THE NETHERLANDS
2 (bases 1 to 573)
Zeeuwen, P.L., Hendriks, W., de Jong, W.W. and Schalkwijk, J.
Identification and sequence analysis of two new members of the
SKALP/elafin and SPAI-2 gene family. Biochemical properties of the
transglutaminase substrate motif and suggestions for a new
nomenclature
J. Biol. Chem. 272 (33), 20471-20478 (1997)
97400522
PUBMED 9252357
REFERENCE 3 (bases 1 to 573)
AUTHORS Molhuizen, H.O., Alkemade, H.A., Zeeuwen, P.L., de Jongh, G.J.,
Wieringa, B. and Schalkwijk, J.
SKALP/elafin: an elastase inhibitor from cultured human
keratinocytes. Purification, cDNA sequence, and evidence for
transglutaminase cross-linking
J. Biol. Chem. 268 (16), 12028-12032 (1993)
93280175

7685029
REFERENCE 4 (bases 1 to 573)
AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
PUBMED 8636131
FEATURES
source
Location/Qualifiers
1..573
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="tongue"
1..405
/gene="bTrappin-2"
<1..405
/gene="bTrappin-2"
/function="transglutaminase substrate, putative proteinase
inhibitor"
/codon_start=1
/protein_id="CAA1184.1"
/db_xref="GI:2764786"
/db_xref="SPTREMBL:O46625"
/translation="QEPVKGDPPVKGDPPVKGDPPVKGDPPVKGRIGGPGLLTKPGSCPVLIRCAM
MNPNNRLRDAQCPCGPKGKCCSGSCGKTCMDPQ"
<1..402
/gene="bTrappin-2"
/product="unnamed"
/function="transglutaminase substrate, putative proteinase
inhibitor"
mat_peptide
BASE COUNT 163 a 158 c 145 g 107 t
ORIGIN

Alignment Scores:
Pred. No.: 1,45e-19 Length: 573
Score: 250.00 Matches: 42
Percent Similarity: 77.42% Conservative: 6
Best Local Similarity: 67.74% Mismatches: 8
Query Match: 76.45% Indels: 6
DB: 4 Gaps: 1

09-833799-13B (1-57) x BOT223216 (1-573)
QY 2 GlnGluProValLys-----GlyProValSerThrLysProGlySer 15
Db 217 CAAGATCCAGTCAAAGGCAAGCGCAATCGAGGTCCACTCTCTCACTAAGCCTGGTCC 276
QY 16 CysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35
Db 277 TGCCCCAGGGTTCGTATCGGTGTCCTATGATGAGAACCCCTTAACCGATGCTCTGAGGAT 336
QY 36 ThrAspCysProGlyLysLysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 337 GCTCAGTGGCCAGGGGTCAAGAAAGTCTGTGAAGGCTCTTGTGGGAAGACCTGTATGGAT 396
QY 56 ProGln 57
Db 397 CCCCAG 402

RESULT 23
LOCUS PIGWAPA
DEFINITION PIG DNA for elafin, complete cds.
ACCESSION D50319
VERSION D50319.1 GI:1228057
KEYWORDS elafin.
SOURCE Sus scrofa DNA, clone lib:EMBL SP6/T7 clone:lambda WAP-1.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (sites)
REFERENCE 1 (sites)

```

AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3693)
AUTHORS Hirose, S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3693)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-924-5824)

FEATURES

source Location/Qualifiers
1..3693
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="lambda WAP-1"
/clone_lib="EMBL SP6/T7"
<1..710
/number=1
<1..628
join(629..710,1541..1962)
/codon_start=1
/product="elafin"
/protein_id="BAA08854.1"
/db_xref="GI:1228058"
/translation="MRSRSLVAVFLIGTIVAAAGRIIRPKGKTKKILALVKG
QGVRGDKQVKGQPYKQDLGKSDPKVKAQLPKGDLGKSDSVKGDPEKAQLPD
KLDPVKAQPAIKRLILITRPGSPRILRLMVPNPNRCLSDAQCPGLKCEGFCG
KAGMDPK"
sig_peptide 629..691
mat_peptide 1777..1959
intron 711..1540
/number=1
1541..1963
/number=2
3'UTR join(1963,2445..2601)
intron 1964..2444
/number=2
repeat_unit 2141..2402
/note="SINE (PRE-1)"
exon 2445..2601
/number=3
polyA_signal 2580..2585
polyA_site 2601
repeat_region 3606..3669
/note="TG repeat"
BASE COUNT 798 a 953 c 1112 g 830 t
ORIGIN

Alignment Scores:
Pred. No.: 1,78e-15 Length: 3693
Score: 221.50 Matches: 39
Percent Similarity: 69.84% Conservative: 5
Best Local Similarity: 61.90% Mismatches: 12
Query Match: 67.74% Indels: 7
DB: 4 Gaps: 1

09-833799-13B (1-57) x PTGWAPA (1-3693)

QY 2 GlnGluProValIysGlyPro-----ValSerThriIysProGly 14
Db 1771 CAAGATCCAGTCAAGGCCCAACCTGCAATCAACGCTCTAATCTTACTTACCAAGCCTGGC 1830
QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsnArgCysLeuIys 34
Db 1831 TCCTGCCCCAGGATCTGATCCGTTGCTTGGTGGTCAATCCCTTACAGGTTGTTGAGT 1890

QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 1891 GATGCTCAGTCCAGGCTCAAGAGTGTCTGTAAGCCTTTTCGGGAGAGCCTGTATG 1950
QY 55 ValProGln 57
Db 1951 GATCCCAAG 1959
RESULT 24
AB003281
LOCUS
DEFINITION Phacochoerus aethiopicus gene for elafin (trappin-2), partial cds.
ACCESSION AB003281
VERSION AB003281.1 GI:4887637
KEYWORDS elafin (trappin-2).
SOURCE Phacochoerus aethiopicus
ORGANISM Phacochoerus aethiopicus
1 (sites) Phacochoerus aethiopicus
REFERENCE 1 (sites) Phacochoerus aethiopicus
AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and
Hirose, S.
JOURNAL Evolution of the trappin multigene family in the Suidae
MEDLINE J. Biochem. 124 (3), 491-502 (1998)
REFERENCE 2 (bases 1 to 1128)
AUTHORS Hirose, S.
JOURNAL Direct Submission
TITLE Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-924-5824)
FEATURES
source Location/Qualifiers
1..1128
/organism="Phacochoerus aethiopicus"
/db_xref="taxon:85517"
<211..632
/codon_start=3
/product="elafin (trappin-2)"
/protein_id="BAA77825.1"
/db_xref="GI:4887638"
/translation="RPKKGKTRIHVLVKQKPVRGEDQVKGQVVKGDILKQGDPPV
KAQLPDKGDLGKEDSVKQDPIKAQLPDKVDPVKAQPAIKRLILITRPGSPRIL
IRCMVNPNNRCLSDAQCPGVKKGCEGFCGKECLNPR"
exon 211..5632
BASE COUNT 296 a 275 c 322 g 235 t
ORIGIN
Alignment Scores:
Pred. No.: 1.46e-15 Length: 1128
Score: 217.50 Matches: 38
Percent Similarity: 68.25% Conservative: 5
Best Local Similarity: 60.32% Mismatches: 13
Query Match: 66.51% Indels: 7
DB: 4 Gaps: 1
09-833799-13B (1-57) x AB003281 (1-1128)
QY 2 GlnGluProValIysGlyPro-----ValSerThriIysProGly 14
Db 441 CAAGATCCAGTCAAGGCCCAACCTGCAATCAACGCTCTAATCTTACTTACCAAGCCTGGC 500
QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsnArgCysLeuIys 34
Db 501 TCCTGCCCCAGGATCTGATCCGTTGATGGTCAATCCCCCTAACAGGTTGTTGAGT 560
QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 561 GATGCTCAGTCCAGGCTCAAGAGTGTCTGTAAGGCTTTTTCGGGAGAGGATGTTGTTG 620
QY 55 ValProGln 57

```

Db      621 AATCCAGG 629
||||:|
RESULT 25
PIGWAPD
LOCUS      PIGWAPD      464 bp      mRNA      linear      MAM 14-APR-2000
DEFINITION Pig mRNA for elafin family member protein, complete cds.
ACCESSION D50322
VERSION    D50322.1 GI:1228063
KEYWORDS   elafin family member protein.
SOURCE     Sus scrofa cDNA to mRNA, clone:WAP-5.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (sites)
            Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
            Tachibana,S. and Hirose,S.
            Accelerated evolution in inhibitor domains of porcine elafin family
            members
            J. Biol. Chem. 271 (12), 7012-7018 (1996)
TITLE      Evolution of the trappin multigene family in the Suidae
JOURNAL    J. Biochem. 124 (3), 491-502 (1998)
MEDLINE    98391820
AUTHORS    Hirose,S.
TITLE      Direct Submission
JOURNAL    Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
            Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
            Midori-ku, Yokohama, Kanagawa 226-8501, Japan
            (E-mail:shiroyee@bio.titech.ac.jp, Tel:045-924-5726,
            Fax:045-924-5824)
FEATURES   Location/Qualifiers
            source
            1..464
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone="WAP-5"
            <1..96
            /number=1
            primer_bind
            1..19
            /note="PCR primer"
            <1..14
            15..446
            CDS
            1..446
            /codon_start=1
            /product="elafin family member protein"
            /protein_id="BAA08857.1"
            /db_xref="GI:1228064"
            /translation="MRSRSFLVVLICGTLVQAGRIIRPKGKTKTLALVKG
            QGPVRGKDQVKGQGPVKQDGLGKSQDPVKAQLPDQDPVKAQPAIKRLLILTKPGSC
            PRILIRCLMVNPNRCLSDAQCPGVKKCCGFCGKDCMDPK"
            15..77
            sig_peptide
            mat_peptide
            261..443
            /product="unnamed"
            97..447
            /number=2
            exon
            447..464
            3'UTR
            448..464
            /number=3
            primer_bind
            complement(448..464)
            /note="PCR primer"
            126 a 111 c 124 g 103 t
            BASE COUNT
            ORIGIN
            Alignment Scores:
            Pred. No.: 7.48e-16 Length: 464
            Score: 216.50 Matches: 38
            Percent Similarity: 68.25% Conservative: 5
            Best Local Similarity: 60.32% Mismatches: 13
            Query Match: 66.21% Indels: 7
            DB: 4 Gaps: 1
09-833799-13B (1-57) x PIGWAPD (1-464)
Qy      2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
        |||:|||||
        Db      255 CAAGTCCAGTGAAGAGCCCAACCTGCAATCAAAAGCTCTAATCTTACTCACAAGCCTGGC 314
        |||:|||||
        Qy      15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsnArgCysLeuLys 34
        |||:|||||
        Db      315 TCCTGCCCCAGGATTCTGATCCGTTGCTTGATGCTCAATCCCTCAACAGGTGTTGAGT 374
        |||:|||||
        Qy      35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
        |||:|||||
        Db      375 GATGCTCAGTGCCAGGGGCAAGAGTCTGTGAAGGCTTTTGGCGGAAGGACTGTATG 434
        |||:|||||
        Qy      55 ValProGln 57
        |||:|||||
        Db      435 GATCCCAAG 443
RESULT 26
D83668
LOCUS      D83668      1034 bp      DNA      linear      MAM 14-APR-2000
DEFINITION Sus scrofa gene for elafin homolog, exon2, partial cds.
ACCESSION D83668
VERSION    D83668.1 GI:2055271
KEYWORDS   elafin homolog; elastase inhibitor.
SOURCE     Sus scrofa DNA, clone_lib:PCR product.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 1034)
            Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
            Tachibana,S. and Hirose,S.
            Accelerated evolution in the inhibitor domain of porcine elafin
            family members
            Journal of Biological Chemistry (1996) In press
JOURNAL    Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and
            Hirose,S.
            Evolution of the trappin multigene family in the Suidae
            J. Biochem. 124 (3), 491-502 (1998)
MEDLINE    98391820
AUTHORS    Hirose,S.
TITLE      Direct Submission
JOURNAL    Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of
            Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
            Midori-ku, Yokohama, Kanagawa 226-8501, Japan
            (E-mail:shiroyee@bio.titech.ac.jp, Tel:045-924-5726,
            Fax:045-924-5824)
COMMENT    On Apr 29, 1997 this sequence version replaced gi:1209380.
            Sequence updated (21-Apr-1997).
FEATURES   Location/Qualifiers
            source
            1..1034
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone_lib="PCR product"
            <1..208
            /number=1
            intron
            1..24
            primer_bind
            /note="PCR primer"
            209..559
            /number=2
            exon
            <209..558
            CDS
            /note="exon2 part"
            /codon_start=3
            /product="elafin homolog"
            /protein_id="BAA12038.1"
            /db_xref="GI:1209381"
            /translation="RPRKGGTKYKTLALVKGQGPVRGKDQVKGQGPVKQDGLGKSQDPV
            KAQLPDQDPVKAQPAIKRLLILTKPGSCPRILIRCLMVNPNRCLSDAQCPGVKK
            CEGFCGKDCMDPK"

```

```

mat_peptide      373..555
/product="elafin homolog"
/function="elastase specific inhibitor"
/notes="Mature protein has a characteristic
disulfide-linked structure called four disulfide core or
WAP motif"
intron           560..>1034
/number=2
primer_bind      1010..>1034
/notes="PCR primer"
BASE COUNT      263 a 262 c 298 g 211 t
ORIGIN
Alignment Scores:
Pred. No.:      1.73e-15      Length:      1034
Score:          216.50      Matches:      38
Percent Similarity: 68.25%      Conservative: 5
Best Local Similarity: 60.32%      Mismatches: 13
Query Match:    66.21%      Indels:      7
DB:             4      Gaps:      1

09-833799-13B (1-57) x DB93668 (1-1034)
Qy 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
Db 367 CAAGATCCAGTCAAGCCCAACCTGCAATCAACAGCTCTAATCTTACTCACCAGGCTGGC 426
Qy 15 SerCysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
Db 427 TCCTGCCCCAGGATTCTCATCGTTGCTTGATGGTCAATCCCTTACAGGTGTTGAGT 486
Qy 35 AspThrAspCysProGlyIleLysLysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 487 GATGCTCAGTCCCGAGGGTCAAGAAGTGTCTGAAGGCTTTTGGGGAAGGACTGTATG 546
Qy 55 ValProGln 57
Db 547 GATCCCAAG 555

RESULT 27
AB003283      AB003283      1169 bp      DNA      linear      MAM 26-MAY-1999
DEFINITION    Pecari tajacu gene for trappin, partial cds.
ACCESSION     AB003283
VERSION       AB003283.1 GI:4887641
KEYWORDS      trappin.
SOURCE        Pecari tajacu DNA.
ORGANISM      Pecari tajacu
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Tayassuidae; Pecari.
1 (sites)
Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and
Hirose,S.
Evolution of the trappin multigene family in the Suidae
J. Biochem. 124 (3), 491-502 (1998)
8391820
REFERENCE      2 (bases 1 to 1169)
AUTHORS        Hirose,S.
TITLE          Direct Submission
JOURNAL        Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences, 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.itech.ac.jp, Tel:045-924-5726,
Fax:045-924-5824)
Location/Qualifiers
1..1169
/organism="Pecari tajacu"
/db_xref="taxon:9829"
<212..675
/codon_start=3
/product="trappin"
/protein_id="BAA77827.1"
/db_xref="GI:4887642"

FEATURES
source
CDS
/translation="RPKGGTGGKHALIKGQDPVRGGPVGQDPVKAQIPDKQDLV
KGDPVKGQDLVKGQDPVKAQIPDKQDPVKGQDPVKGQDPVKAQIPDKQDLV
LPKPGKPGKPMIKIRCALFPNRLTDAAGCGARKCCIGSCGKALNVR"
212..675
exon           305 a 289 c 330 g 245 t
ORIGIN
Alignment Scores:
Pred. No.:      1.14e-13      Length:      1169
Score:          201.00      Matches:      35
Percent Similarity: 65.08%      Conservative: 6
Best Local Similarity: 55.56%      Mismatches: 14
Query Match:    61.47%      Indels:      8
DB:             4      Gaps:      1

09-833799-13B (1-57) x AB003283 (1-1169)
Qy 2 GlnGluProValLysGlyProValSerThr-----LysPro 13
Db 478 CAAGATCCAGTCAAGCCCAACCTTGCAGTCAGAGCTCTGGTCTTACCCCGCGTAAGCCT 537
Qy 14 GlySerCysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCysLeu 33
Db 538 GGCTTCGCGCCCATGATTAAAGATCCGTTGTGCCCTGTTCAATCTCTCTTAAAGGTTTG 597
Qy 34 LysAspThrAspCysProGlyIleLysLysCysGluGlySerCysGlyMetAlaCys 53
Db 598 ACCGATGCTGGGTGCCAGGCGCCAGGAAGTCTGCATAGGCTCTTGGCGGAAGGCGCTGT 657
Qy 54 PheValPro 56
Db 658 TTGAATCCT 666

RESULT 28
MAM223215      MAM223215      270 bp      mRNA      linear      PRI 15-DEC-1998
LOCUS         Macaca mulatta mRNA for putative sTrappin-2 protein, partial.
DEFINITION    Macaca mulatta mRNA for putative sTrappin-2 protein, partial.
ACCESSION     AJ223215
VERSION       AJ223215.1 GI:2764783
KEYWORDS      proteinase inhibitor; putative; sTrappin-2 gene; transglutaminase
substrate.
SOURCE        rhesus monkey.
ORGANISM      Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
1 (bases 1 to 270)
Zeeuwen,P.L.J.M.
Direct Submission
JOURNAL        Submitted (06-JAN-1998) Zeeuwen P.L.J.M., Dermatology, University
Hospital Nijmegen, PO BOX 9101, 6500 HB Nijmegen, THE NETHERLANDS
2 (bases 1 to 270)
Zeeuwen,P.L., Hendriks,W., de Jong,W.W. and Schalkwijk,J.
Identification and sequence analysis of two new members of the
SKALP/elafin and SPAL-2 gene family. Biochemical properties of the
transglutaminase substrate motif and suggestions for a new
nomenclature
J. Biol. Chem. 272 (33), 20471-20478 (1997)
97400522
MEDLINE        97400522
PUBMED         9252357
REFERENCE      3 (bases 1 to 270)
AUTHORS        Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
Tachibana,S. and Hirose,S.
Accelerated evolution in inhibitor domains of porcine elafin family
members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
96215132
MEDLINE        96215132
PUBMED         8636131
REFERENCE      4 (bases 1 to 270)
AUTHORS        Molhuizen,H.O., Alkenade,H.A., Zeeuwen,P.L., de Jongh,G.J.,
Wieringa,B. and Schalkwijk,J.
SKALP/elafin: an elastase inhibitor from cultured human
keratinocytes. Purification, cDNA sequence, and evidence for

```


OS	Sus sp.	(pig)	COMMENT
PN	JP 1994045098-A/1		
PD	22-FEB-1994		
PF	29-JUL-1992	JP 199202286	
PI	KUROKI ATSUSHI, KUMATA MANABU, TACHIBANA SHINRO	PC	
CK	K73/00, A61K37/64, A61K49/00, C12N15/12		
CC	strandedness: Double;		

```

CC anti-sense: No;
CC *source: tissue_type=Duodenum;
CC *source: library=lamdba gt10 cdna library;
CC *source: clone=PD-57G;
FH Key Location/Qualifiers
FH CDS
FH FT 1..570
FH FT /product='SPAI'
FH FT 385..567
FH FT /product='SPAI'
FH FT misc_feature 70..567
FH FT /note='The region corresponds to h
FH FT molecular SPAI'
FH FT 3'UTR 571..737.
FH FT Location/Qualifiers
FH FT 1..737
FH FT /organism="Sus sp."
FH FT /db_xref="taxon:9826"
FH FT 205 a 192 c 182 g 158 t
FH FT BASE COUNT
FH FT FEATURES
FH FT source

```

```

Alignment Scores:
Pred. No.:      1.95e-11          Length:      737
Score:         179.50           Matches:     31
Percent Similarity: 61.90%       Conservative: 8
Best Local Similarity: 49.21%    Mismatches:  17
Query Match:   54.89%           Indels:      7
DB:            6                Gaps:        1

09-833799-13B (1-57) x E06782 (1-737)

Qy      2  GlnGluProValLysGlyPro-----ValSerThrLysPro
      |||:::||::||::||::||::||::||::||::||::||::||::||
Db      379 CAAGTCCAGTCGAAGGCCAACCTGCAGTCCGAGTCGATTCCTCTCTCAAGCGTT

Qy      15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      439 CACTGCCCTAGGATCTTTTTGGTTGCCGCTGAGCATCCCTCTAACAAGTGTTGG

Qy      35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetalaCys
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      499 GATTATGACTGCCAGGGGTCAAGAAGTGCTGTGAAGGCTTTTGGCGGAAGGATTGT

Qy      55 ValProGln 57
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      559 TATCCCAAG 567

```

RESULT 32	
DMY245	
LOCUS	789 bp mRNA linear MAM
DEFINITION	Sus scrofa mRNA for preproSPAI-2, complete cds.
ACCESSION	D83667 D17753
VERSION	G1:1304046
KEYWORDS	APAI-2; preproSPAI-2; proSPAI-2; SPAI-2.
SOURCE	Sus scrofa cell_line:NMS14 cDNA to mRNA, clone_lib:lambd
ORGANISM	Sus scrofa
	Eukaryota; Metazoa; Chordata; Vertebrata; Eut
	Mammalia; Eutheria; Cetartiodactyla; Suina Suidae; Sus

```

REFERENCE 1 (sites)
AUTHORS Kuroki,J., Hosoya,T., Itakura,M., Hirose,S., Tamechika,I.,
        Yoshimoto,T., Ghoneim,M.A., Nara,K., Kato,A., Suzuki,Y.,
        Furukawa,M. and Tachibana,S.
TITLE Cloning, characterization, and tissue distribution of porcine SPAl,
        a protein with a transglutaminase substrate domain and the WAP
        motif
JOURNAL J. Biol. Chem. 270 (38), 22428-22433 (1995)
MEDLINE 95403443
REFERENCE 2 (bases 1 to 789)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of
        Technology, Department of Biological Sciences; Okayama, Meguro-ku,
        Tokyo 152, Japan (Tel:81-3-3726-1111, Fax:81-3-3729-0335)
COMMENT On or before Mar 17, 1999 this sequence version replaced
        gi:2078450. gi:1054611, gi:2077950.
        D17753:Submitted (27-Sep-1993) to DDBJ by:Shigehisa Hirose.
FEATURES
    source
        1. .789
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /cell_line="NM514"
            /clone_lib="lambda gt10"
            61..624
            /codon_start=1
            /product="preproSPAI-2"
            /protein_id="BAA12037.1"
            /db_xref="GI:1304181"
            /translation="MRSRSFLVLAVFLICETLVAQRDLDRIRGPKGGQDPVEGQDD
            EGPGPVKEILDIGDPVKGDQDPVKGDQDPVKGDQDPVKQDLPVKSDPVKAEPLDIGD
            VKGHEPVEGQDPVNAQLPDQDPVKQDPVAPGPRFLSKRGHCPRILFRCLSPNSN
            KCWRDYDCPGVKKCCGFCGKDCLYPK"
            61..123
            sig_peptide
            mat_peptide
            124..621
            /product="proSPAI-2"
            439..621
            mat_peptide
            /product="SPAI-2"
            722
            polyA_site
            polyA_signal
            761..766
            BASE COUNT 201 a 215 c 197 g 176 t
            ORIGIN
            Alignment Scores:
            Pred. No.: 2,09e-11 Length: 789
            Score: 179.50 Matches: 31
            Percent Similarity: 61.90% Conservative: 8
            Best Local Similarity: 49.21% Mismatches: 17
            Query Match: 54.89% Indels: 7
            DB: 4 Gaps: 1
09-833799-13B (1-57) x DMV245 (1-789)
QY 2 GinGluproVallysglyPro-----ValSerThrlypsProGly 14
Db 433 CAAGATCCAGTCAAGCCCAACCTGCAGTCCAGGTGCGATTCTCTCTAAGCGTGC 492
QY 15 SerCysProilleleLeulleArgCysAlaMetLeuAsnProAsnArgCysLeuLys 34
Db 493 CACTGCCCTAGGATCTTTTTCGTCGCCGCTGAGCAATCCCTCTAACAGGTGTTGGAGA 552
QY 35 AspThrAspCysProGlyllelylsysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 553 GATTATGACTGTCCAGGGGTCAAGAGTCTGTGAAGGCTTTTCGCGGAAGGATTGTTTG 612
QY 55 ValProGln 57
Db 613 TATCCCAAG 621
RESULT 33
PIGWAPB PIGWAPB 3782 bp DNA linear MAM 14-APR-2000
LOCUS
DEFINITION Pig DNA for SPAl-2, complete cds.

```

```

ACCESSION D50320
VERSION D50320.1
KEYWORDS GI:1228059
SOURCE SPAl-2.
ORGANISM Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-2.
REFERENCE 1 (sites)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
        Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
        Tachibana,S. and Hirose,S.
        Accelerated evolution in inhibitor domains of porcine elafin family
        members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3782)
AUTHORS Hirose,S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3782)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
        Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
        Midori-ku, Yokohama, Kanagawa 226-8501, Japan
        (E-mail:shirose@bio.citech.ac.jp, Tel:045-924-5726,
        Fax:045-924-5824)
FEATURES
    source
        1. .3782
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone_lib="lambda WAP-2"
            /clone_lib="EMBL SP6/T7"
            <1..710
            /number=1
            /evidence=experimental
            <1..628
            join(629..710,1552..2033)
            /codon_start=1
            /evidence=experimental
            /product="SPAI-2"
            /protein_id="BAA08855.1"
            /db_xref="GI:1228060"
            /translation="MRSRSFLVLAVFLICETLVAQRDLDRIRGPKGGQDPVEGQDD
            EGPGPVKEILDIGDPVKGDQDPVKGDQDPVKGDQDPVKQDLPVKSDPVKAEPLDIGD
            VKGHEPVEGQDPVNAQLPDQDPVKQDPVAPGPRFLSKRGHCPRILFRCLSPNSN
            KCWRDYDCPGVKKCCGFCGKDCLYPK"
            629..691
            sig_peptide
            1848..2030
            mat_peptide
            /product="SPAI-2"
            /evidence=experimental
            711..1551
            intron
            /number=1
            /evidence=experimental
            1552..2034
            exon
            /number=2
            /evidence=experimental
            join(2034,2545..2701)
            2035..2544
            3'UTR
            /number=2
            /evidence=experimental
            2240..2502
            repeat_unit
            /note="SINE (PRE-1)"
            2545..2701
            exon
            /number=3
            /evidence=experimental
            2680..2685
            polyA_signal
            polyA_site
            2701
            repeat_region
            /evidence=experimental
            3671..3742
            /note="TG repeat"
            BASE COUNT 823 a 972 c 1149 g 838 t
            ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.08e-10 Length: 3782
Score: 179.50 Matches: 31
Percent Similarity: 61.90% Conservative: 8
Best Local Similarity: 49.21% Mismatches: 17
Query Match: 54.89% Indels: 7
DB: 4 Gaps: 1

09-833799-13B (1-57) x PIGWAPB (1-3782)
Qy 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
Db 1842 CAAGTCAGTCAAGAACCCACCTGCAGTCCAGTTCCTCTCTAAGCGTGGC 1901
Qy 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
Db 1902 CACTGCCTAGGATCTTTTCGTGTCGCTGAGCATCCCTCAACAAGTGTGGAGA 1961
Qy 35 AspThrAspCysProGlyIleLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 1962 GATTATGACTGTCAGGGGTCAAGAGTGTGTGAAGGCTTTTTCGGGAAGGATTGTTG 2021
Qy 55 ValProGln 57
Db 2022 TATCCCAAG 2030

RESULT 34
PIGWAPB
LOCUS PIGWAPB 578 bp mRNA linear MAM 14-APR-2000
DEFINITION Pig mRNA for elafin family member protein, complete cds.
ACCESSION D50323
VERSION D50323.1 GI:1228065
KEYWORDS elafin family member protein.
SOURCE Sus scrofa cDNA to mRNA, clone: WAP-4.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
J Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE
AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and
Hirose, S.
TITLE Evolution of the trappin multigene family in the Suidae
J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE
AUTHORS Hirose, S.
TITLE Direct Submission
Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences, 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirosebio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-924-5824)
FEATURES
source
1..578
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="WAP-4"
<1..96
/number=1
1..19
/notes="PCR primer"
<1..14
15..560
/codon_start=1
/product="elafin family member protein"
/protein_id="BAA08858.1"
/db_xref="GI:1864016"

exon
primer_bind
5'UTR
CDS

/translation="MRSRSLVLVAVFLICETLVAORPEKIRGKGGQDPVGGQDQD
KGGPVKEVLIDIGDQLVRKQDPVKGDQLVKSDQPVKAEPLDIDGQDVKGH
PVGGQDPVNAQLPKVDQDPVKAQPPVQGRLLHYKGLCPWIFLRCPLPKPPNCKWRDS
HCPGVKMKCCGFCGNECYSYR"
15..77
sig_peptide
mat_peptide
375..557
/product="unnamed"
97..561
exon
number=2
561..578
562..578
number=3
primer_bind
complement(562..578)
/notes="PCR primer"
BASE COUNT 169 a 141 c 138 g 130 t
ORIGIN
Alignment Scores:
Pred. No.: 2.55e-11 Length: 578
Score: 177.50 Matches: 30
Percent Similarity: 67.24% Conservative: 9
Best Local Similarity: 51.72% Mismatches: 18
Query Match: 54.28% Indels: 1
DB: 4 Gaps: 1

09-833799-13B (1-57) x PIGWAPB (1-578)
Qy 1 AlaGlnGluProValLysGlyPro---ValSerThrLysProGlySerCysProIleIle 19
Db 384 GCCCAACCTCCAGTCCAAGGTCGATTCTCACTATAAGCTGCGCTCTGCCCTTGGATT 443
Qy 20 LeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysPro 39
Db 444 TTTTACGTTGCGCGCTGCCCAACCTCTCAACAATGTTGGAGAGATTCTCACTGCCCA 503
Qy 40 GlyIleLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 504 GGGGTCATCAAGTCTGTGAAGGCTTTTTCGGGAATGATGTTCTGATCCCCGA 557

RESULT 35
AF178426
LOCUS AF178426 490 bp mRNA linear ROD 29-AUG-1999
DEFINITION Rattus norvegicus secretory leukocyte protease inhibitor mRNA,
complete cds.
ACCESSION AF178426
VERSION AF178426.1 GI:5802679
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS 1 (bases 1 to 490)
Song, X., Zeng, L., Jin, W., Thompson, J., Mizel, D.E., Lei, K.,
Billingham, R.C., Poole, A.R. and Wahl, S.M.
TITLE Secretory leukocyte protease inhibitor suppresses the inflammation
and joint damage of bacterial cell wall-induced arthritis
J. Exp. Med. 190 (4), 535-542 (1999)
MEDLINE 99380355
PUBMED 10449524
REFERENCE
AUTHORS 2 (bases 1 to 490)
Song, X.-Y. and Zeng, L.
TITLE Direct Submission
Submitted (17-AUG-1999) Oral infection & Immunity Branch, National
Institute of Dental & Craniofacial research, 9000 Rockville Pike,
Bethesda, MD 20892
FEATURES
Location/Qualifiers
1..490
/organism="Rattus norvegicus"
/strain="Lewis"
/db_xref="taxon:10116"
9..404
/notes="slpi"
CDS

```

```

/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAD51758.1"
/db_xref="GI:5802680"
/translation="MKSGLEPLMVLALGLAPWSEGGKNDKIKGACPAKPAQC
LKLEPEGGTDEWCEPKQRKQCDTGFKCLNPVPIRGVKKKPGRCVKFGKCLMLNP
FNKQNDQCGDKYKCCGCMGKVCPLPV"
BASE COUNT 118 a 121 c 131 g 120 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1,178-10 Length: 490
Score: 171.00 Matches: 28
Percent Similarity: 60.38% Conservative: 4
Best Local Similarity: 52.83% Mismatches: 21
Query Match: 52.29% Indels: 0
DB: 10 Gaps: 0

```

09-833799-13B (1-57) x AF178426 (1-490)

```

QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 CCATTCGTGGACCAAGTGAAGAAGACCTGGAGGTGGTCAAAATTTCAAGGAAATGT 299
QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 CTGATGCTTAACCCCTCCCAATAGTGCAGAAATGACGGCCAGTGTGATGTCACAAATACAAA 359
QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 TGTGTGAGGGCATGTGTGGGAAAGTCTGCCTTCCCCCA 398

```

```

RESULT 36
AB011010
LOCUS

```

```

DEFINITION Bos taurus gene for Trappin-6, partial cds.
ACCESSION AB011010
VERSION AB011010.1 GI:3132273
KEYWORDS Trappin-6.
SOURCE Bos taurus DNA.

```

```

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.

```

```

TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 495)
AUTHORS Hirose,S.

```

```

TITLE Direct Submission
JOURNAL Submitted (10-FEB-1998) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shiros@bio.titech.ac.jp, Tel:+81-45-924-5726, Fax:+81-45-924-5824)

```

```

FEATURES
source
1..495
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..36
/number=1
primer_bind 1..24
/note="PCR primer site"
CDS
<37..329
/codon_start=3
/product="Trappin-6"
/protein_id="BAA28148.1"
/db_xref="GI:3132274"
/translation="SPKGGNVVFNKGKPVNGSPDKGDPVKGDVPVKGDVVVAQD
RAGLPFRKGLCPVRIRHCNLWNFPNQCRDHAHCFGAKKCCGFCGTCWNPR"

```

```

exon 37..>329
primer_bind 472..495
/note="PCR primer site"

```

```

BASE COUNT 134 a 99 c 163 g 99 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1,548-10 Length: 495
Score: 170.00 Matches: 30
Percent Similarity: 56.06% Conservative: 7
Best Local Similarity: 45.45% Mismatches: 19
Query Match: 51.99% Indels: 10
DB: 4 Gaps: 1

```

09-833799-13B (1-57) x AB011010 (1-495)

```

QY 2 GlnGluProValLysGlyPro-----ValSerThr 11
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 CAAGACCCAGTCAAGAGCAAGATGTAGTCAGACAGCGCGGACTTCATTCC 188
QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArg 31
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 AAGCGTGGCTTATGCCCCAGGGTTCGGATCCACTGCAACTTGTGGAAATCCCTAAC 248
QY 32 CysLeuLysAspThrAspCysProGlyIleLeuLysLysCysGluGlySerCysGlyMet 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 249 TGTGGAGAGATGCTCACTGCCAGGGGCCAAGAGTGTGTGAAGCTTTTGTGGGAA 308
QY 52 AlaCysPheValProGln 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 309 ACCTGTATGAATCCCGG 326

```

```

RESULT 37
AB003284
LOCUS

```

```

DEFINITION Hippopotamus amphibius gene for trappin, partial cds.
ACCESSION AB003284
VERSION AB003284.1 GI:4887643
KEYWORDS trappin.
SOURCE Hippopotamus amphibius DNA.

```

```

ORGANISM
Hippopotamus amphibius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.

```

```

REFERENCE
1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.

```

```

TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 734)
AUTHORS Hirose,S.

```

```

TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shiros@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)

```

```

FEATURES
source
1..734
/organism="Hippopotamus amphibius"
/db_xref="taxon:9833"
<1..545
/codon_start=3
/product="trappin"
/protein_id="BAA77828.1"
/db_xref="GI:4887644"
/translation="HPKQGTGKGNVLFKGQDPVKGDVPVKGOYPVKGDVPVKGDVPVKGD
PVKGDVPVKGOYPVKGDVPVKGDVPVKGDVPVKGDVPVKGDVPVKGDVPVKGD
CPRNKNCCPASCCKICKLFP"

```

```

BASE COUNT 245 a 158 c 198 g 133 t
ORIGIN

```

Alignment Scores:

Pred. No.: 6.64e-10 734
 Score: 166.00 28
 Percent Similarity: 61.54% 4
 Best Local Similarity: 53.85% 20
 Query Match: 50.76% 0
 DB: 4 0
 Gaps: 0

09-833799-13B (1-57) x AB003284 (1-734)

Qy 2 GlnGluProValSerThrLysProGlySerCysProIleLeuLeuLeu 21

Db 375 CAAGATCCAGTCAAGTCCATTCCTGCTCGGAAGTGTCTGCCCAAGATTGGATT 434

Qy 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLeu 41

Db 435 GAGTGTCTCCACATTAAATCCCAAAACGGGTGTGAGAGATGCTCAGTCCCAAGGAAC 494

Qy 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53

Db 495 AAGATTGCTGTCGGCCCTCTTGGGGAAGATCTGT 530

RESULT 38

PIGALP

LOCUS PIGALP 600 bp mRNA linear MAM 07-MAR-1995
 DEFINITION Porcine antileukoproteinase mRNA, complete cds.
 ACCESSION M57446
 VERSION M57446.1 GI:164319
 KEYWORDS Porcine uterus. cDNA to mRNA, clone pALP.
 SOURCE Sus scrofa

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 600)
 AUTHORS Farmer,S.J., Fliss,A.E. and Simmen,R.C.
 TITLE Complementary DNA cloning and regulation of expression of the
 messenger RNA encoding a pregnancy-associated porcine uterine
 protein related to human antileukoproteinase
 Mol. Endocrinol. 4 (8), 1095-1104 (1990)

JOURNAL

MEDLINE

PUBMED

FEATURES

Location/Qualifiers

1..600
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone="pALP"
 /sex="female"
 /tissue_type="uterus"
 1..600
 /gene="ALP"
 44..391
 /gene="ALP"
 /codon_start=1
 /product="antileukoproteinase"
 /protein_id="AAA63446.1"
 /db_xref="GI:164320"

gene

CDS

polyA_site
 KCCRDTCALKCLNPAITNPVKRPGKCPVYVYGCMMNPPNCKTDSQCLGLDKCK
 SMCKVKCLTPVKA
 600

BASE COUNT 146 a 156 c 151 g 147 t

ORIGIN

Alignment Scores:

Pred. No.: 6.98e-10 Length: 600
 Score: 165.00 Matches: 25
 Percent Similarity: 59.62% Conservative: 6
 Best Local Similarity: 48.08% Mismatches: 21
 Query Match: 50.46% Indels: 0
 DB: 4 0
 Gaps: 0

09-833799-13B (1-57) x PIGALP (1-600)

Qy 5 VallysGlyProValSerThrLysProGlySerCysProIleLeuLeuLeuLeu 24
 Db 224 ATCAGAACCCAGTTAAGTGAAGCCTGGGAAGTGTCCAGTGGTCTATGGCCAGTGTATG 283
 Qy 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLeuLysCys 44
 Db 284 ATGCTCAACCCCCCAATCACTGCAAGACACAGACAGCCAGTGCCTGGGTGACTTAAATGC 343

Qy 45 CysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 344 TGCAAGACATGTGGGAAAGTGTGCTCACCCT 379

RESULT 39

MMU88093

LOCUS MMU88093 671 bp mRNA linear ROD 19-APR-1997
 DEFINITION Mus musculus secretory leukocyte protease inhibitor mRNA, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 671)
 AUTHORS Zitnik,R.J., Zhang,J., Kashem,M.A., Kohno,T., Lyons,D.E.,
 Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
 TITLE The cloning and characterization of a murine secretory leukocyte
 protease inhibitor cDNA

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..671
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="ras-transformed 3T3 cells"
 13..408
 /function="serine protease inhibitor"
 /codon_start=1
 /product="secretory leukocyte protease inhibitor"
 /protein_id="AAC53140.1"
 /db_xref="GI:1945383"
 /translation="MKSGGLLPFTVLALGILAPWTEGGKNDAIKIGACPAKPAQC
 LKLEKPCQRTDWECPGKQRCQDAGSKCNFVPIRKPVWRKPGRCVKTQRCMNLN
 PNVQRDQCDGKYKCEGICGKVLPPM"

BASE COUNT 157 a 179 c 177 g 158 t

ORIGIN

Alignment Scores:

Pred. No.: 2.24e-09 Length: 671
 Score: 161.00 Matches: 28
 Percent Similarity: 58.49% Conservative: 3
 Best Local Similarity: 52.83% Mismatches: 22
 Query Match: 49.24% Indels: 0
 DB: 10 0
 Gaps: 0

09-833799-13B (1-57) x MMU88093 (1-671)

Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuLeuLeuLeu 23

Db 244 CCATTTCGCAACCACTGTGGAGGAAGCTGGAGTGGCTCAAACTCAGCAAGATGT 303

Qy 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLeuLys 43

Db 304 ATGATGCTTAACCCCTCCCAATGCTGCCAGGAGGCGGCGAGTGTGACGCAATACAAG 363

Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 |||||
 Db 364 TGTGTGAGGGTATATGTGGGAAAGTCTGCTGCCCG 402
 |||||
 RESULT 40
 MMU94341
 LOCUS
 DEFINITION Mus musculus secretory leukoprotease inhibitor mRNA, complete cds.
 ACCESSION U94341
 VERSION U94341.1 GI:1945450
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Abe, T., Tomimaga, Y., Kikuchi, T., Watanabe, A., Satoh, K., Watanabe, Y.
 and Nukiwa, T.
 TITLE Bacterial pneumonia causes augmented expression of the secretory
 leukoprotease inhibitor gene in the murine lung
 JOURNAL Am. J. Respir. Crit. Care Med. 156 (4 Pt 1), 1235-1240 (1997)
 MEDLINE 98011992
 PUBMED 9351627
 REFERENCE
 AUTHORS 2 (bases 1 to 682)
 Abe, T., Tomimaga, Y., Kikuchi, T., Watanabe, A., Satoh, K., Watanabe, Y.
 and Nukiwa, T.
 TITLE Direct Submission
 JOURNAL Submitted (18-MAR-1997) Department of Respiratory Oncology and
 Molecular Medicine, Institute of Development, Aging and Cancer,
 Tohoku University, 4-1 Seiryomachi, Aoba-ku, Sendai 980-77, Japan
 FEATURES
 source
 1. .682
 /organism="Mus musculus"
 /strain="B6CBA [C57BL/6 x CBA]"
 /db_xref="taxon:10090"
 /tissue_type="lung"
 14. .409
 /codon_start=1
 /product="secretory leukoprotease inhibitor"
 /protein_id="AAC51394.1"
 /db_xref="GI:1945451"
 /translation="MKSCGLPFTVLLALGILAPWTVEGGKNDKIGACPAKPPAQ
 LKLEKPCORTDWECPGKQCCODAGSKCVNPVIRKPVWRKPCVCVKTCQARCMMLNP
 PNVQRDQGDGKYKCEGICGKVCPLPM"

BASE COUNT 163 a 179 c 181 g 159 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.27e-09 Length: 682
 Score: 161.00 Matches: 28
 Percent Similarity: 58.49% Conservativeness: 3
 Best Local Similarity: 52.83% Mismatches: 22
 Query Match: 49.24% Indels: 0
 DB: 10 Gaps: 0
 09-833799-13B (1-57) x MMU94341 (1-682)

Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
 |||||
 Db 245 CCCATTCGCAACACGCTGTGGAGGAGCTGTGGAGGTCGTCACAACTCAGGCAAGATGT 304
 |||||
 Qy 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
 |||||
 Db 305 ATGATGCTTAACCTCCCAATGTCTGCAGAGGACGGGAGTGTGACGCAATACAAG 364
 |||||
 Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 |||||
 Db 365 TGTGTGAGGGTATATGTGGGAAAGTCTGCTGCCCG 403
 |||||

RESULT 41
 BC028509
 LOCUS 894 bp mRNA linear ROD 07-AUG-2002

DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT

US
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Marcello Bento Soares, Ph.D.
 CDNA Library Preparation: Soares Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK Plate: 67 Row: j Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6755573.
 FEATURES
 source
 1. .894
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:41142 IMAGE:1513866"
 /tissue_type="Mammary gland, lactating mouse"
 /clone_lib="Soares_mammary_gland_NMLMG"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac"
 236. .631
 /codon_start=1
 /product="secretory leukocyte protease inhibitor"
 /protein_id="AAH28509.1"
 /db_xref="GI:20306996"
 /translation="MKSCGLPFTVLLALGILAPWTVEGGKNDKIGACPAKPPAQ
 LKLEKPCORTDWECPGKQCCODAGSKCVNPVIRKPVWRKPCVCVKTCQARCMMLNP
 PNVQRDQGDGKYKCEGICGKVCPLPM"

BASE COUNT 205 a 242 c 233 g 214 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.02e-09 Length: 894
 Score: 161.00 Matches: 28
 Percent Similarity: 58.49% Conservativeness: 3
 Best Local Similarity: 52.83% Mismatches: 22
 Query Match: 49.24% Indels: 0
 DB: 10 Gaps: 0
 09-833799-13B (1-57) x BC028509 (1-894)

Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
 |||||
 Db 467 CCCATTCGCAACACGCTGTGGAGGAGCTGTGGAGGTCGTCACAACTCAGGCAAGATGT 526
 |||||

```

QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLeuLysLys 43
Db 527 ATGATGCTTAACCTCCCAATGCTGCAGAGGACGGGAGTGTGACGCAATACAAG 586

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 587 TGCTGTGAGGATATATGTGGGAAAGTCTGCTGCCCCCG 625

RESULT 42
MMU73004
LOCUS MMU73004 1123 bp mRNA linear ROD 27-FEB-1997
DEFINITION Mus musculus secretory leukocyte protease inhibitor mRNA, complete
cds.
ACCESSION U73004
VERSION U73004.1 GI:1763262
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1123)
AUTHORS Jin, F.Y., Nathan, C., Radzioch, D. and Ding, A.
TITLE Secretory leukocyte protease inhibitor: a macrophage product
induced by and antagonistic to bacterial lipopolysaccharide
CELL 88 (3), 417-426 (1997)
JOURNAL Cell 88 (3), 417-426 (1997)
MEDLINE 97191310
PubMed 9039268
REFERENCE 2 (bases 1 to 1123)
AUTHORS Ding, A., Jin, F.-Y. and Nathan, C.F.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-1996) Medicine, Cornell University Medical
College, 1300 York Ave, Box 57, New York, NY 10021, USA
FEATURES
source
location/Qualifiers
1..1123
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="RAW 264.7"
/cell_type="macrophage"
447..842
/function="serine protease inhibitor"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAC53047.1"
/db_xref="GI:1763263"
/translation="MKSCGLLPFTLVLLALGILAPWTVEGGKNDKIKGACPAKKPAQC
LKLEKPCORTDWECPGKQRCODACGSKCNVPVPIKPVWKPGRVCVKQARCMNLNP
PNCQRGQCDGKYKCEGICGKVCLEPPM"
BASE COUNT 286 a 297 c 288 g 252 t
ORIGIN
Alignment Scores:
Pred. No.: 3.84e-09 Length: 1123
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x MMU73004 (1-1123)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuLysArgCys 23
Db 678 CCCATTCCGCAACCAAGTGTGGAGGAAGCGCTGGAGGTGCGTCAAAACATCAGGCAAGATGT 737

QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLeuLysLys 43
Db 738 ATGATGCTTAACCTCCCAATGCTGCAGAGGACGGGAGTGTGACGCAATACAAG 797

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 798 TGCTGTGAGGATATATGTGGGAAAGTCTGCTGCCCCCG 836

```

```

RESULT 43
AF151982
LOCUS AF151982 667 bp mRNA linear ROD 01-JUN-1999
DEFINITION Rattus norvegicus secretory leukocyte protease inhibitor (SLPI)
mRNA, complete cds.
ACCESSION AF151982
VERSION AF151982.1 GI:4929548
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 667)
AUTHORS Chen, D.H., Xu, X.P., Bagchi, M.K. and Bagchi, I.C.
TITLE Molecular cloning and spatio-temporal expression of rat secretory
leukocyte protease inhibitor (SLPI) in the uterus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 667)
AUTHORS Chen, D.H., Xu, X.P., Bagchi, M.K. and Bagchi, I.C.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1999) Center for Biomedical Research, The
Population Council, 1230 York Avenue, New York, NY 10021, USA
FEATURES
source
location/Qualifiers
1..667
/organism="Rattus norvegicus"
/strain="SP"
/db_xref="taxon:10116"
1..667
/gene="SLPI"
13..405
/gene="SLPI"
/function="serine protease inhibitor"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAD34035.1"
/db_xref="GI:4929549"
/translation="MKSSGLFPFLMVLALGLVLPWSVEGGKNDKIKGACPAKKPAQC
LKREKPCSTDWGCPGKQRCQDTCGFKCLNPPVPIRGVKKPKGKCLAFQKCLMLNPP
NKQNDGQCDGKYKCEGICGKVCLEPPV"
BASE COUNT 176 a 159 c 161 g 171 t
ORIGIN
Alignment Scores:
Pred. No.: 2.53e-09 Length: 667
Score: 160.50 Matches: 28
Percent Similarity: 60.38% Conservative: 4
Best Local Similarity: 52.83% Mismatches: 20
Query Match: 49.08% Indels: 1
DB: 10 Gaps: 1

09-833799-13B (1-57) x AF151982 (1-667)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuLysArgCys 23
Db 244 CCCATTCCGTCGACCAAGTGAAG---AAGCTGGAGGTGCTCAAAATTTCAAGGAAATGT 300

QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLeuLysLys 43
Db 301 CTGATGCTTTAACCTCCCAATGAAGTGCCAGATGACGGCCAGTGTGATGGCAATACAAA 360

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 361 TGTGTGAGGATGTGTGGGAAAGTCTGCTGCCCTTCCCCCA 399

RESULT 44
PIGWAPC
LOCUS PIGWAPC 3670 bp DNA linear MAM 14-APR-2000
DEFINITION Pig DNA for elafin family member protein, complete cds.
ACCESSION D50321
VERSION D50321.1 GI:1228061
KEYWORDS elafin family member protein.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-3.

```


Db	1343	GGCCTGTGAATCTTCTTCTCAAGTGTGGAGGAAGCCTGGGAGGTGGCTGCTCAAACTCAG	1402
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	1403	GCAAGATGTATGATGCTTAACCCCTCCCAATGCTCTGCCAGAGGACGGCAGTGTGACGGC	1462
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	1463	AAATACAAAGTGCTGTGAGGGTATATGTGGAAAGTCTGCCTGCCCCCG	1510
RESULT	48		
AL590429/c			
LOCUS	AL590429	190669 bp	DNA linear ROD 17-NOV-2001
DEFINITION	Mouse DNA sequence from clone RP23-462016 on chromosome 2, complete sequence.		
ACCESSION	AL590429		
VERSION	AL590429.9	GI:17017748	
KEYWORDS	HTG.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 190669)		
TITLE	Kay, M.		
JOURNAL	Direct Submission		
COMMENT	Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Nov 20, 2001 this sequence version replaced gi:15020991. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30). An attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-462016 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6 This sequence is the entire insert of clone RP23-462016 The true left end of clone RP23-140D14 is at 187423 in this sequence. The true right end of clone RP23-346D16 is at 20063 in this sequence.		
FEATURES	Location/Qualifiers		
source	1..190669 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="2" /clone="RP23-462016" /clone_lib="RPCI-23" 104714..104881 /note="Sequence from uni-directional dGTP big dye terminator reads only"		
misc_feature	140443..140460 /note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."		
misc_feature	140460 /note="Random repeat. Forced join. Gap size estimated to be approximately 900bp by restriction digest data."		
BASE COUNT	55113 a	43425 c	42104 g 50027 t
ORIGIN			


```

VERSION      AX328399.1  GI:18098360
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Rastelli, L. and Smithson, G.
TITLE        Method of detecting inflammatory lung disorders
JOURNAL      Patent: WO 0190421-A 1 29-NOV-2001;
              Curagen Corporation (US)
FEATURES
source
1.594
/organism="Homo sapiens"
/db xref="taxon:9606"
BASE COUNT   132 a 155 g 151 t
ORIGIN
Alignment Scores:
Pred. No.:   2.19e-07      Length: 594
Score:       143.00       Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73%      Indels: 2
DB:          Gaps: 1

09-833799-13B (1-57) x AX328399 (1-594)
Qy  3  GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db  238  GATCTGTGTGACACCCCAACCAAGAGGAGGCTGGGAAGTCCCAAGTCACTTAT 297
Qy  21  IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db  298  GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGCCAGTCAAGCGT 357
Qy  41  IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db  358  GACTTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 405

RESULT 59
AX334507
LOCUS      AX334507          594 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION Sequence 5885 from Patent WO0194629.
ACCESSION  AX335376
VERSION     AX335376.1  GI:18126095
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
              Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE        Cancer gene determination and therapeutic screening using signature
              gene sets
JOURNAL      Patent: WO 0194629-A 5885 13-DEC-2001;
              Avalon Pharmaceuticals (US)
FEATURES
source
1.594
/organism="Homo sapiens"
/db xref="taxon:9606"
BASE COUNT   132 a 155 c 155 g 151 t
ORIGIN
Alignment Scores:
Pred. No.:   2.19e-07      Length: 594
Score:       143.00       Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73%      Indels: 2
DB:          Gaps: 1

09-833799-13B (1-57) x AX334507 (1-594)
Qy  3  GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db  238  GATCTGTGTGACACCCCAACCAAGAGGAGGCTGGGAAGTCCCAAGTCACTTAT 297
Qy  21  IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db  298  GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGCCAGTCAAGCGT 357
Qy  41  IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db  358  GACTTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 405

RESULT 59
AX334507
LOCUS      AX334507          594 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION Sequence 5016 from Patent WO0194629.
ACCESSION  AX334507
VERSION     AX334507.1  GI:18125226
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
              Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE        Cancer gene determination and therapeutic screening using signature
              gene sets
JOURNAL      Patent: WO 0194629-A 5016 13-DEC-2001;
              Avalon Pharmaceuticals (US)
FEATURES
source
1.594
/organism="Homo sapiens"
/db xref="taxon:9606"
BASE COUNT   132 a 155 c 155 g 151 t
ORIGIN
Alignment Scores:
Pred. No.:   2.19e-07      Length: 594
Score:       143.00       Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73%      Indels: 2
DB:          Gaps: 1

```


repeat_region /note="match: GSS: Em:AQ760198"
 13052..13153
 /note="MIR repeat: matches 33..137 of consensus"
 misc_feature 13131..13392
 /note="match: GSS: Em:AQ081155"
 repeat_region 13218..13397
 /note="45 copies 4 mer gaaa 66% conserved"
 13928..14216
 /note="AluX repeat: matches 1..310 of consensus"
 repeat_region 14372..14593
 /note="L1ME2 repeat: matches 5935..6162 of consensus"
 repeat_region 14895..15168
 /note="AluX repeat: matches 5..295 of consensus"
 *misc_feature complement(14950..15573)
 /note="match: GSS: Em:AQ241185"
 misc_feature 15703..16159
 /note="match: GSS: Em:AQ017550"
 repeat_region 16175..16190
 /note="L1M4 repeat: matches 4145..4159 of consensus"
 repeat_region 16191..16500
 /note="AluSg repeat: matches 2..310 of consensus"
 repeat_region 16501..16608
 /note="L1M4 repeat: matches 4027..4145 of consensus"
 repeat_region 16659..16839
 /note="MER97b repeat: matches 2..179 of consensus"
 repeat_region 16847..17348
 /note="L1M4 repeat: matches 3473..4025 of consensus"
 repeat_region 17616..19408
 /note="L1MB7 repeat: matches 4367..6167 of consensus"
 repeat_region 19499..19670
 /note="86 copies 2 mer at 79% conserved"
 repeat_region 19502..19669
 /note="42 copies 4 mer tata 79% conserved"
 repeat_region 19678..21535
 /note="L1M4 repeat: matches 67..1979 of consensus"
 repeat_region 21547..21682
 /note="Tigger4 (Zombi) repeat: matches 1..137 of consensus"
 repeat_region 21690..22537
 /note="Tigger4 (Zombi) repeat: matches 1872..2730 of consensus"
 repeat_region 22669..22996
 /note="L1M4 repeat: matches 3623..3953 of consensus"
 repeat_region 23636..23896
 /note="L1M2 repeat: matches 1943..2572 of consensus"
 repeat_region 23825..26024
 /note="L1 repeat: matches 2121..4337 of consensus"
 misc_feature 25994..26445
 /note="match: GSS: Em:AQ470847"
 repeat_region 26025..26160
 /note="LTR13 repeat: matches 796..936 of consensus"
 repeat_region 26161..26235
 /note="L1 repeat: matches 5112..5188 of consensus"
 misc_feature 26204..26681
 /note="match: GSS: Em:AQ624124"
 *misc_feature complement(27205..27381)
 /note="match: STS: Em:G06121"
 gene complement(27206..29528)
 /gene="SLPI"
 mRNA complement(join(27206..27386,27969..28118,28542..28700,29426..29528))
 /gene="SLPI"
 /product="dJ30012.2 (secretory leukocyte protease inhibitor (antileukoproteinas))"
 Em:X04503 Em:M57446 Em:U94341 Em:AF151982 Em:AF178426 Em:AF114471 Em:J31216
 match: ESTs: Em:R71834 Em:AA683520 Em:AA316675 Em:AI742512 Em:AA572950 Em:AA564454 Em:AI862145 Em:AI540954 Em:AI638119 Em:AI222907 Em:AI743345 Em:AW081599 Em:AW264225 Em:AA460433 Em:AI885550 Em:AW238407 Em:AI564623 Em:AA541595 Em:AA132992 Em:H65117 Em:AA397776 Em:AA993549 Em:AI37093 Em:AA165295 Em:AW190157 Em:AA557989 Em:AA164414 Em:H65171 Em:N27733 Em:AW103765

Em:AA460434 Em:R49913 Em:AA026192 Em:AW391622 Em:T28664 Em:AA026099 Em:AA026497 Em:AA485776 Em:AI247078 Em:AI276322 Em:AI367954 Em:AI826892"
 /evidence=not experimental
 complement(27206)
 /gene="SLPI"
 complement(27226..27231)
 /gene="SLPI"
 /gene="SLPI"
 27229..27315
 /note="L2 repeat: matches 2627..2710 of consensus"
 complement(27230..27235)
 /gene="SLPI"
 complement(join(27382..27386,27969..28118,28542..28700,29426..29510))
 /gene="SLPI"
 /note="match: proteins: Tr:O44397 Sw:P03973 Tr:Q9WUQ4 Sw:P97430 Tr:Q9R028 Tr:O95959 Sw:P09412 Sw:P22298 Tr:Q9XS44"
 /codon_start=1
 /evidence=not experimental
 /product="dJ30012.2 (secretory leukocyte protease

polyA_site
 polyA_signal
 repeat_region
 polyA_signal
 CDS

Alignment Scores:
 Pred. No.: 2,95e-05 Length: 63796
 Score: 143.00 Matches: 25
 Percent Similarity: 52.83% Conservative: 3
 Best Local Similarity: 47.17% Mismatches: 25
 Query Match: 43.73% Indels: 0
 DB: 9 Gaps: 0

09-833799-13B (1-57) x HS30012 (1-63796)
 QY 4 ProValLysGlyProValSerThrLysProGlySerCysProLleLeuLleA19Cys 23
 Db 28134 CCTACTCTACCAGCAAGAGGAGCCCTGGGAAGTCCAGTACTTATGGCCATGT 28075
 QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLleLysLys 43
 Db 28074 TTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTCAAGCGTGAAG 28015
 QY 44 CysCysGlyGlySerGlyMetAlaCysPheValPro 56
 Db 28014 TGTTGATGGCATGTGTGGAAATCTCGCTTCCCT 27976
 RESULT 65
 HSSLIPG
 LOCUS HSSLIPG 2657 bp DNA linear PRI 30-MAR-1995
 DEFINITION Human SLPI gene for secretory leukocyte protease inhibitor.
 ACCESSION X04502
 VERSION X04502.1 GI:36485
 KEYWORDS elastase inhibitor; protease inhibitor; secretory leucocyte protease inhibitor; trypsin inhibitor.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Stetler,G., Brewer,M.T. and Thompson,R.C.
 Isolation and sequence of a human gene encoding a potent inhibitor of leukocyte proteases
 Nucleic Acids Res. 14 (20), 7883-7896 (1986)
 JOURNAL MEDLINE 87040761
 PUBMED 3640338
 REFERENCE 2 (bases 1 to 2657)
 AUTHORS Rogaeve,E.I., Keryanov,S.A. and Malyako,Y.K.
 TITLE Dinucleotide repeat polymorphisms at the P1, HBE1 and MYH7 loci Hum. Mol. Genet. 1 (4), 285 (1992)
 JOURNAL MEDLINE 93265040
 PUBMED 1363870
 FEATURES
 Location/Qualifiers
 1..2657
 /organism="Homo sapiens"
 /db_xref="taxon:9606"


```

source      1. .321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene        1. .321
/gene="ALP-246"
CDS         1. .321
/partial
/gene="ALP-246"
/codon_start=1
/transl_table=11
/protein_id="CAA00748.1"
/db_xref="GI:413351"
/translation="SGKSFKAGVCPPKSAOCLRYKKPECDWQCPGKRCPCDPTCG
IKCLDPVDTNPTRRKPGKCPVTGQCLLLNPNFCEMDGQCKRDLKCCMGCKGSCV
SPVKA"
BASE COUNT 72 a 84 c 92 g 73 t
ORIGIN

Alignment Scores:
Pred. No.:      4,248-07      Length:      321
Score:          138.00      Matches:      26
Percent Similarity: 55.36%      Conservative: 5
Best Local Similarity: 46.43%      Mismatches: 23
Query Match:    42.20%      Indels:      2
DB:              6      Gaps:      1

09-833799-13B (1-57) x A08112 (1-321)
QY  3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db  145 GATCCGGTTGACACCCGACCGGCGTCTGTAACCCGGAAGTCCCGGTACCGAA 204
QY  21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db  205 GGTCAGTGCCTGCTGTAACCCGCTAACTTCTCGAGATGATGGCCAGTGCAAAACGA 264
QY  41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db  265 GATCTGAATGCTGATGGTATGTCGGGTAAAGCTCGTTAGCCCG 312
RESULT 68
A08107
LOCUS      A08107      321 bp      DNA      linear      PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 231.
ACCESSION  A08107
VERSION    A08107.1 GI:413350
KEYWORDS   synthetic construct.
SOURCE     synthetic construct.
ORGANISM   artificial sequences.
REFERENCE  1 (bases 1 to 321)
AUTHORS    Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE      Serine protease inhibitor proteins, medicaments containing them,
           DNA sequences coding for these proteins and methods for producing
           these proteins, medicaments and DNA sequences
           Patent: EP 037335-A 29 20-JUN-1990;
           Gruenthal GmbH
JOURNAL
FEATURES
source
Location/Qualifiers
1. .321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene
1. .321
/gene="ALP-231"
CDS
1. .321
/partial
/gene="ALP-231"
/codon_start=1
/transl_table=11
/protein_id="CAA00743.1"
/db_xref="GI:413351"
/translation="SGKSFKAGVCPPKSAOCLRYKKPECDWQCPGKRCPCDPTCG
IKCLDPVDTNPTRRKPGKCPVTGQCLLLNPNFCEMDGQCKRDLKCCMGCKGSCV
SPVKA"
BASE COUNT 72 a 84 c 92 g 73 t
ORIGIN

Alignment Scores:
Pred. No.:      5,518-07      Length:      321
Score:          137.00      Matches:      26
Percent Similarity: 55.36%      Conservative: 5
Best Local Similarity: 46.43%      Mismatches: 23
Query Match:    41.90%      Indels:      2
DB:              6      Gaps:      1

09-833799-13B (1-57) x A08107 (1-321)
QY  3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db  145 GATCCGGTTGACACCCGACCGGCGTCTGTAACCCGGAAGTCCCGGTACCTAC 204
QY  21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db  205 GGTCAGTGCCTGCTGTAACCCGCTAACTTCTCGAGATGATGGCCAGTGCAAAACGA 264
QY  41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db  265 GATCTGAATGCTGCTGCGTCTGTCGGGTAAAGCTCGTTAGCCCG 312
RESULT 69
A08108
LOCUS      A08108      321 bp      DNA      linear      PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 232.
ACCESSION  A08108
VERSION    A08108.1 GI:413352
KEYWORDS   synthetic construct.
SOURCE     synthetic construct.
ORGANISM   artificial sequences.
REFERENCE  1 (bases 1 to 321)
AUTHORS    Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE      Serine protease inhibitor proteins, medicaments containing them,
           DNA sequences coding for these proteins and methods for producing
           these proteins, medicaments and DNA sequences
           Patent: EP 037335-A 30 20-JUN-1990;
           Gruenthal GmbH
JOURNAL
FEATURES
source
Location/Qualifiers
1. .321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene
1. .321
/gene="ALP-232"
CDS
1. .321
/partial
/gene="ALP-232"
/codon_start=1
/transl_table=11
/protein_id="CAA00744.1"
/db_xref="GI:413353"
/translation="SGKSFKAGVCPPKSAOCLRYKKPECDWQCPGKRCPCDPTCG
IKCLDPVDTNPTRRKPGKCPVTGQCLLLNPNFCELDGQCKRDLKCCGLGCKGSCV
SPVKA"
BASE COUNT 69 a 89 c 89 g 74 t
ORIGIN

Alignment Scores:
Pred. No.:      5,518-07      Length:      321
Score:          137.00      Matches:      26
Percent Similarity: 55.36%      Conservative: 5
Best Local Similarity: 46.43%      Mismatches: 23
Query Match:    41.90%      Indels:      2
DB:              6      Gaps:      1

09-833799-13B (1-57) x A08108 (1-321)
QY  3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db  145 GATCCGGTTGACACCCGACCGGCGTCTGTAACCCGGAAGTCCCGGTACCTAC 204
QY  21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db  205 GGTCAGTGCCTGCTGTAACCCGCTAACTTCTCGAGATGATGGCCAGTGCAAAACGA 264
QY  41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db  265 GATCTGAATGCTGCTGCGTCTGTCGGGTAAAGCTCGTTAGCCCG 312

```

```

Db 145 GATCGGTTGACACCCGAAACCGGCGTCTGTAACCCCGGAAGTCCCGGTTACCTAC 204
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGTCAGTGCATCCTGCTGAACCCCGCTTAACCTTCGCGAGCTCGATGGCCAGTGCACAAACGA 264
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAATGCTGCTGGGCTGCTGGGTAAAGCTGCTAGCCCG 312

RESULT 70
A08110
LOCUS A08110 321 bp DNA linear PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 237.
ACCESSION A08110
VERSION A08110.1 GI:413356
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 321)
AUTHORS Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE Serine protease inhibitor proteins, medicaments containing them,
DNA sequences coding for these proteins and methods for producing
these proteins, medicaments and DNA sequences
JOURNAL Patent: EP 0373335-A 32 20-JUN-1990;
Gruenthal GmbH
FEATURES
source
Location/Qualifiers
1..321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene 1..321
/length=321
CDS 1..321
partial
/gene="ALP-237"
/codon_start=1
/transl_table=11
/protein_id="CAA00746.1"
/db_xref="GI:413357"
/translation="SGKSFKAGVCPPKKSAQCLRYKKPEQSDWQCPGKRCPPDTCG
IKCLDPVDTNPTRRKPKGKCPVITGQCLLNPPNFCELDGCKRDLKCCCLGLGKSCV
SPVKA"
BASE COUNT 68 a 89 c 89 g 75 t
ORIGIN

Alignment Scores:
Pred. No.: 5,51e-07 Length: 321
Score: 137.00 Matches: 26
Percent Similarity: 55.36% Conservative: 5
Best Local Similarity: 46.43% Mismatches: 23
Query Match: 41.90% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x A08110 (1-321)
Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GATCGGTTGACACCCGAAACCGGCGTCTGTAACCCCGGAAGTCCCGGTTACCTAC 204
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGTCAGTGCCTGCTGCTGAACCCGCTTAACCTTCGCGAGCTCGATGGCCAGTGCACAAACGA 264
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAATGCTGCTGGGCTGCTGGGTAAAGCTGCTAGCCCG 312

RESULT 71
A31082
LOCUS A31082 74 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI5 from patent EP0402068.
ACCESSION A31082

```

```

VERSION A31082.1 GI:1249296
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 74)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 22 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
Location/Qualifiers
1..74
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 19 a 14 c 18 g 23 t
ORIGIN

Alignment Scores:
Pred. No.: 2e-07 Length: 74
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.28% Indels: 0
DB: 6 Gaps: 0

09-833799-13B (1-57) x A31082 (1-74)
Qy 36 ThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 3 ACTGATTGTCAGGATCAAAAAGTCTGTGAAGTTCTCTGCGGTATGCTTTGTTGTT 62
Qy 56 ProGln 57
Db 63 CCACAA 68

RESULT 72
A08109
LOCUS A08109 321 bp DNA linear PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 236.
ACCESSION A08109
VERSION A08109.1 GI:413354
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 321)
AUTHORS Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE Serine protease inhibitor proteins, medicaments containing them,
DNA sequences coding for these proteins and methods for producing
these proteins, medicaments and DNA sequences
JOURNAL Patent: EP 0373335-A 31 20-JUN-1990;
Gruenthal GmbH
FEATURES
source
Location/Qualifiers
1..321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene 1..321
/length=321
CDS 1..321
partial
/gene="ALP-236"
/codon_start=1
/transl_table=11
/protein_id="CAA00745.1"
/db_xref="GI:413355"
/translation="SGKSFKAGVCPPKKSAQCLRYKKPEQSDWQCPGKRCPPDTCG
IKCLDPVDTNPTRRKPKGKCPVITGQCLLNPPNFCELDGCKRDLKCCCLGLGKSCV
SPVKA"
BASE COUNT 69 a 88 c 91 g 73 t
ORIGIN

Alignment Scores:

```

```

Pred. No.: 9.3e-07 Length: 321
Score: 135.00 Matches: 26
Percent Similarity: 55.36% Conservative: 5
Best Local Similarity: 46.43% Mismatches: 23
Query Match: 41.28% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x A08109 (1-321)

Qy 3 GluProValysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GATCCGGTTGACACCCGACCGACGCTGCTAAACCCGGGAAGTCCCGGTTACCGAA 204
Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGTCAAGTGCCTGCTGCTGAACCGCGCTAACTCTCGAGCTGATGGCCAGTCAACGA 264
Qy 41 IleLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAATGCTGCTGGGTCTGTGCGGTAAAGCTGCTTAGCCCG 312

RESULT 73
AB003285
LOCUS AB003285 1054 bp DNA linear MAM 26-MAY-1999
DEFINITION Sus scrofa gene for elafin homolog, partial cds.
ACCESSION AB003285
VERSION AB003285.1 GI:4887645
KEYWORDS elafin homolog.
SOURCE Sus scrofa DNA.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and
Hirose, S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1054)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel:045-924-5726,
Fax:045-924-5824)
FEATURES
source Location/Qualifiers
1..1054
/organism="Sus scrofa"
/db_xref="taxon:9823"
<211..557
/codon_start=3
/product="elafin homolog"
/protein_id="BAA77829.1"
/db_xref="GI:4887646"
/translation="GLKGQDDPVAGQAGEGPVKVEILDIGDLVKSDPVEQDDPV
RAQLPDKVDQDPVKAQPAVQGVFPFVKLGFCEIRCLLNPNRCLIDAQCPGFQKCC
RVCGVKCADPR"
exon 211..557
BASE COUNT 263 a 262 c 307 g 222 t
ORIGIN
Alignment Scores:
Pred. No.: 3.24e-06 Length: 1054
Score: 135.00 Matches: 32
Percent Similarity: 49.40% Conservative: 9
Best Local Similarity: 38.55% Mismatches: 14
Query Match: 41.28% Indels: 28
DB: 4 Gaps: 5

09-833799-13B (1-57) x AB003285 (1-1054)
Qy 1 AlaGlnGluProValysGly----- 7

```

```

...:|||||:|||||:|||||:
Db 312 AGTCAAGATCCAGTTGAGGGTCAAGATCCAGTCCAGGCCCACTTCCAGACAAAGTACAA 371
Qy 8 ---ProValSerThrLysPro-----GlySer 15
Db 372 GATCCAGTCAAAAGCCCAACCTGCAGTCCAAGGTGATTCTCTCTCTCTCTCTCTCTCT 431
Qy 16 CysProIleLeuIleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAsp 35
Db 432 TGCCCCAGGATTGATCCCGTTGCAGACTGCTCAAT---CTAACAGGTGTTTATAGAT 488
Qy 36 ThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet---AlaCysPhe 54
Db 489 GCTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 545
Qy 55 ValProGln 57
Db 546 GATCCCCGG 554

RESULT 74
A08106
LOCUS A08106 Synthetic ALP-gene 240.
DEFINITION A08106
ACCESSION A08106
VERSION A08106.1 GI:413348
KEYWORDS synthetic construct.
SOURCE artificial sequences.
ORGANISM 1 (bases 1 to 321)
REFERENCE Heindel-Wieland, R., Ammann, J., Steffens, G.J. and Flohe, L.
AUTHORS Serine protease inhibitor proteins, medicaments containing them,
TITLE DNA sequences coding for these proteins and methods for producing
these proteins, medicaments and DNA sequences
JOURNAL Patent: EP 0373335-A 28 20-JUN-1990;
Gruenthal GmbH
FEATURES
source Location/Qualifiers
1..321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene 1..321
/genes="ALP-240"
CDS 1..321
/partial
/gene="ALP-240"
/codon_start=1
/transl_table=1
/protein_id="CAA00742.1"
/db_xref="GI:413349"
/translation="SGKSPKAGVCPKPKSAOCLRYKKPEQSDWOCPGKKRCCPDTC
IKCLDPVDTFNPTRPKGKCPVTYGCILLNRPNCMDGQCKRDLKCCMGKSCV
SPVKA"
BASE COUNT 72 a 83 c 92 g 74 t
ORIGIN
Alignment Scores:
Pred. No.: 2.65e-06 Length: 321
Score: 131.00 Matches: 25
Percent Similarity: 53.57% Conservative: 5
Best Local Similarity: 44.64% Mismatches: 24
Query Match: 40.06% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x A08106 (1-321)
Qy 3 GluProValysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GATCCGGTTGACACCCCGACCGACGCTGCTAAACCCGGGAAGTCCCGGTTACCGAA 204
Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGTCAAGTGCCTGCTGCTGAACCGCGCTAACTCTCGAGCTGATGGCCAGTCAACGA 264

```


GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 19:00:51 ; Search time 223 Seconds
(without alignments)
575.623 Million cell updates/sec

Title: 09-833799-13b

Perfect score: 327

Sequence: 1 aqepvkgpvstkpqscpiil.....cpqikckcecgscgmactvpg 57

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO spoil/US0983799/runat 12022003 120041 23624/app query.fasta_1.199
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0983799 @CGN 1 1 396 @runat 12022003 120041 23624 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	327	100.0	206	11	AAQ06819
C 2	327	100.0	434	21	AAC68808
C 3	327	100.0	480	24	ABK45842
4	327	100.0	504	11	AAQ06820
5	327	100.0	571	24	ABQ58818
6	327	100.0	2309	24	ABK83819
7	327	100.0	2309	24	ABL65816
8	327	100.0	2309	24	ABL66488
9	324	99.1	177	15	AAQ56674
10	324	99.1	8598	19	AAV28852
11	323	98.8	177	15	AAQ56675
12	323	98.8	177	15	AAQ56676
13	322	98.5	321	13	AAQ28757
14	179.5	54.9	737	15	AAQ44862
15	161	49.2	396	24	ABQ73674
C 16	161	49.2	396	24	ABQ73677
17	161	49.2	684	20	AAI18516
18	161	49.2	691	20	AAI18514
19	161	49.2	1114	20	AAI18515
20	146	44.6	180	21	AAC97526
21	146	44.6	183	20	AAI16273
22	146	44.6	183	22	AAI67584
23	143	43.7	194	10	AAAN90354
24	143	43.7	321	24	ABK8016
25	143	43.7	324	7	AAAN60463
26	143	43.7	324	7	AAAN60464
27	143	43.7	324	20	AAI16194
28	143	43.7	324	20	AAI16236
29	143	43.7	324	21	AAC97527
30	143	43.7	324	21	AAC97528
31	143	43.7	324	21	AAC97529
32	143	43.7	324	22	AAI67488
33	143	43.7	324	22	AAI67489
34	143	43.7	399	22	ABA8108
35	143	43.7	420	20	AAI16272
36	143	43.7	436	21	AAC97624
37	143	43.7	436	22	AAI67583
38	143	43.7	460	20	AAI16271
39	143	43.7	460	21	AAC97622
40	143	43.7	460	22	AAI67582
41	143	43.7	498	8	AAAN70929
42	143	43.7	565	8	AAAN70930
43	143	43.7	565	22	AAF32043
44	143	43.7	594	24	ABL66679
45	143	43.7	594	24	ABL67548
46	143	43.7	594	24	AAI17499
47	143	43.7	599	20	AAZ77546
48	143	43.7	636	22	AAH57451
49	143	43.7	1525	24	ABK88022
50	143	43.7	1525	24	ABK88025
51	141	43.1	2274	7	AAAN60465
52	141	43.1	2274	21	AAC97530
53	141	43.1	2274	22	AAI67491
54	139	42.5	180	15	AAQ45442
55	139	42.5	180	21	AAZ88483
56	123.5	37.8	1010	21	AAA37665
57	123.5	37.8	1013	24	AAD38698
58	122	37.3	321	21	AAC97529
59	121.5	37.2	1411	21	AAZ51673
60	118.5	36.2	558	22	AAH78211
61	118.5	36.2	732	22	AAH78210
62	118.5	36.2	5690	22	AAH07251
63	118.5	36.2	5690	23	ABL98799
64	116	35.5	872	7	AAAN60469
65	116	35.5	920	7	AAAN60468
66	112	34.3	478	20	AAAX80909
67	112	34.3	762	21	AAA99904

Description

68 112 34.3 762 21 AA265078 Membrane-bound pro
 69 112 34.3 762 22 AAF44224 Human PRO844 (UNQ5
 70 107 32.7 411 24 AB644114 Human benign prost
 71 107 32.7 411 24 AB66677 Lung cancer relate
 72 107 32.7 411 24 AB67461 Thyroid cancer relat
 73 107 32.7 411 24 AB67811 Ovary cancer relat
 74 105 32.1 292 21 AA70066 Human ovarian carc
 75 105 32.1 292 24 ABN72960 Ovarian carcinoma
 76 103 31.5 4536 23 AB109753 Drosophila melanog
 77 103 31.5 8853 23 AB109752 Drosophila melanog
 78 101 30.9 1796 23 AB10651 Drosophila melanog
 79 101 30.9 7930 23 AB10650 Drosophila melanog
 80 98 30.0 705 22 AAF59010 Mouse epididymis-s
 81 98 30.0 724 22 AAD10127 Mouse major epidid
 82 98 30.0 724 24 AB135060 Murine cDNA isolat
 83 98 30.0 767 24 AAD37774 Extended sequence
 84 98 30.0 8577 23 AB128667 Drosophila melanog
 85 98 30.0 9295 23 AB102495 Drosophila melanog
 86 98 30.0 12294 23 AB128666 Drosophila melanog
 87 98 30.0 21314 23 AB102494 Drosophila melanog
 88 97 29.7 120 21 AAC97577 Insert C used in p
 89 97 29.7 120 21 AAC97620 DNA insert used in
 90 97 29.7 120 22 AA167539 Serine protease in
 91 97 29.7 583 12 AAQ13127 HE4 epididymis-spe
 92 97 29.7 583 19 AAV64622 Human HE4 cDNA. H
 93 97 29.7 583 22 AAD02714 Human epididymal p
 94 97 29.7 583 24 ABL68022 Ovary cancer relat
 95 96.5 29.5 308 24 ABN25109 Human OREX polynuc
 96 96.5 29.5 401 22 AA54133 Murine transport a
 97 95.5 29.2 500 21 AAC93641 Cat flea head and
 98 95 29.1 469 22 ABA42966 Human breast cell
 99 95 29.1 469 22 ABA23163 Probe #1629 for ge
 100 95 29.1 469 22 AAK01653 Human brain expres

ALIGNMENTS

RESULT 1
 AAQ06819
 ID AAQ06819 standard; DNA; 206 BP.
 XX AAQ06819;
 XX
 XX 06-MAR-1991 (first entry)
 DT
 DE Sequence encoding human leukocyte elastase inhibiting peptide.
 DE
 DE
 KW Emphysema; respiratory distress syndrome; atherosclerosis;
 KW arthritis; cystic fibrosis; leukaemia.
 XX
 XX Homo sapiens.
 XX
 XX
 XX Key Location/Qualifiers
 FT CDS 30...203
 FT /*tag= a
 FT /label= Elastase inhibitor
 XX
 XX EP402068-A.
 XX
 XX
 XX 12-DEC-1990.
 PD
 XX
 XX 04-JUN-1990; 90EP-0306037.
 PF
 XX 02-NOV-1989; 89GB-0024717.
 PR 09-JUN-1989; 89GB-0013346.
 PR 09-JUN-1989; 89GB-0013349.
 PR 25-SEP-1989; 89GB-0021613.
 XX
 XX (ICIL) IMPERIAL CHEM INDS PLC.
 PA
 XX Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
 PI
 XX WPI, 1990-370081/50.
 DR

DR P-PSDB; AAR08217.
 XX Human leukocyte elastase inhibitor - is genetically engineered
 PT polypeptide for treatment of inflammatory, pulmonary and skin
 PT conditions
 XX
 PS Disclosure; Fig 13; 45pp; English.
 XX
 CC Gene product has inhibitory activity against human leukocyte
 CC elastase (HLE) and may be isolated from psoriatic scales. The
 CC sequence may be expressed from a plasmid transformed expression
 CC system and may be useful in the prevention of tissue damage
 CC associated with emphysema, adult respiratory distress syndrome,
 CC psoriasis and bullous dermatoses. Other treatable conditions
 CC include atherosclerosis, cystic fibrosis, bronchitis and acute
 CC non-lymphoblastic leukaemia.
 CC Abs raised to the polypeptide may be used in detection.
 XX
 SQ Sequence 206 BP; 51 A; 46 C; 44 G; 65 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.47e-29 Length: 206
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 11 Gaps: 0
 09-833799-13B (1-57) x AAQ06819 (1-206)
 QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 30 GCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTGTCTATTATCTT 89
 QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 90 ATTCTGTCGGCTATGTTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAGT 149
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 150 ATCAAAAAGTGTGTGAAGGTTCTCGCGTATGCGTGTGTTTTCGTTCCACAA 200
 RESULT 2
 AAC68808/c
 ID AAC68808 standard; cDNA; 434 BP.
 XX
 XX AAC68808;
 XX
 XX 21-FEB-2001 (first entry)
 DT
 DE Human head/neck tumour related protein partial coding sequence #5.
 DE
 DE
 KW Head tumour; neck tumour; lung cancer; vaccine; cancer therapy; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200065053-A2.
 PN
 XX 02-NOV-2000.
 PD
 XX
 XX 21-APR-2000; 2000WO-US10687.
 PF
 XX 23-APR-1999; 99US-0130906.
 PR 20-APR-2000; 2000US-0533870.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Wang T, Dillon DC;
 PI
 XX WPI; 2000-687345/67.
 DR
 XX Novel polypeptides comprising immunogenic portion of head and neck
 PT tumour protein useful for treating, diagnosing and monitoring cancer

PT such as head, neck and lung cancer -
 XX Claim 3; Page 71; 77pp; English.
 XX
 CC The present invention relates to a number of nucleic acid sequences which
 CC encode proteins associated with head, neck and lung tumours. These
 CC tumours are often not diagnosed until they have spread, and, of those who
 CC survive, most must endure alterations in facial and neck appearance as
 CC well as changes in speech, sight, smell, chewing, swallowing and taste
 CC perception. The coding sequences given, and the proteins they encode, can
 CC be used in the diagnosis, treatment and vaccination against cancer,
 CC particularly papillary and follicular carcinomas, papillary tumours,
 CC follicular adenoma, parathyroid hyperplasia, parotid cancer, lip cancer,
 CC squamous cell cancer of the tongue, oral tongue cancers and larynx
 CC cancer.
 XX
 SQ Sequence 434 BP; 103 A; 103 C; 133 G; 94 T; 1 other;
 Alignment Scores:
 Pred. No.: 6.15e-29 Length: 434
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 09-833799-13B (1-57) x AAC68808 (1-434)
 QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 DB 318 GCGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCCTGGCTGCCCATTTATCTTG 259
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 258 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAGACTGACTGCCCAGGA 199
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 198 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 148
 RESULT 3
 ID ABK45842/c
 XX ABK45842 standard; cDNA; 480 BP.
 AC ABK45842;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE cDNA encoding colon tumour protein, SEQ ID No 1393.
 XX
 KW Human; colon tumour; vaccine; colon cancer; immunogenic;
 KW immunotherapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200212328-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 31-JUL-2001; 2001WO-US24218.
 XX
 PR 03-AUG-2000; 2000US-223283P.
 PR 28-MAR-2001; 2001US-279763P.
 PR 29-JUN-2001; 2001US-302051P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI King GE, Meagher MJ, Xu J, Secrist H;
 XX
 DR WPI; 2002-241739/29.
 XX
 PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
 PT for diagnosing, preventing, and treating colon cancer, and as markers

PT for the progression of cancer -
 XX Claim 1; SEQ ID No 1393; 147pp; English.
 XX
 CC The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer.
 CC ABK4450-ABK46237 represent coding sequences of human colon tumour
 CC proteins of the invention.
 CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
 CC for this patent did not form part of the printed specification but was
 CC supplied by the European Patent Office.
 XX
 SQ Sequence 480 BP; 113 A; 105 C; 139 G; 120 T; 3 other;
 Alignment Scores:
 Pred. No.: 6.96e-29 Length: 480
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 09-833799-13B (1-57) x ABK45842 (1-480)
 QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 DB 369 GCGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCCTGGCTGCCCATTTATCTTG 310
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 309 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAGACTGACTGCCCAGGA 250
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 249 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 199
 RESULT 4
 ID AAQ06820
 XX AAQ06820 standard; cDNA; 504 BP.
 AC AAQ06820;
 XX
 DT 06-MAR-1991 (first entry)
 XX
 DE Sequence encoding human leukocyte elastase inhibiting peptide.
 XX
 KW HLE; Emphysema; respiratory distress syndrome; atherosclerosis;
 KW arthritis; cystic fibrosis; leukaemia; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..157
 FT /*tag= b
 FT /label= Upstream in-frame coding sequence
 CDS 121..328
 FT /*tag= a
 FT /label= Elastase inhibitor
 FT intron 144..180
 FT /*tag= c
 XX
 PN EP402068-A.
 XX
 PD 12-DEC-1990.
 XX
 PF 04-JUN-1990; 90EP-0306037.

```

XX 02-NOV-1989; 89GB-0024717.
XX 09-JUN-1989; 89GB-0013346.
XX 09-JUN-1989; 89GB-0013349.
XX 25-SEP-1989; 89GB-0021613.
XX (ICIL) IMPERIAL CHEM INDS PLC.
XX
XX Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
XX
XX WPI; 1990-370081/50.
XX P-PSDB; AAR08217.
XX
XX Human leukocyte elastase inhibitor - is genetically engineered
XX polypeptide for treatment of inflammatory, pulmonary and skin
XX conditions
XX
XX Disclosure; Fig 16; 45pp; English.
XX
XX Gene product has inhibitory activity against human leukocyte
XX elastase (HLE) and may be isolated from psoriatic scales. The
XX sequence may be expressed from a plasmid transformed expression
XX system and may be useful in the prevention of tissue damage
XX associated with emphysema, adult respiratory distress syndrome,
XX psoriasis and bullous dermatoses. Other treatable conditions
XX include atherosclerosis, cystic fibrosis, bronchitis and acute
XX non-lymphoblastic leukaemia.
XX Abs raised to the polypeptide may be used in detection.
XX
XX Sequence 504 BP; 110 A; 152 C; 120 G; 122 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7.39e-29 Length: 504
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
09-833799-13B (1-57) x AAQ06820 (1-504)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 158 GCGCAAGAGCCAGTCAAGGTCCTCCACTAAGCTGCTGCTGCCCATTTATCTTG 217
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 218 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGACTGCCACGA 277
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 278 ATCAAGAAAGTGTGTGAAGGCTCTTGCGGGATGGCTGTTTCGTTCCTCCAG 328
RESULT 5
ABQ58818
ID ABQ58818 standard; cDNA; 571 BP.
XX
XX ABQ58818;
XX
XX 02-AUG-2002 (first entry)
XX
XX Human colon cancer related nucleotide sequence SEQ ID NO:2513.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200229086-A2.
XX
XX 11-APR-2002.
XX
XX '02-OCT-2001; 2001WO-US30732.
XX
02-OCT-2000; 2000US-237271P.
XX
XX (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX Thiagalingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
XX tissues useful for determining the presence of colon cancer in a cell
XX or tissue type, and in antisense therapy -
XX
XX Claim 1; Fig 1; 796pp; English.
XX
XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX used in antisense therapy. An antibody immunoreactive with a polypeptide
XX encoded by (I) is useful for detecting cancer in a patient sample, and
XX for detecting the presence or absence of a polynucleotide encoded by a
XX nucleic acid which hybridises to (I) in a cell. A probe/primer derived
XX from (I) can be used for determining the presence of a nucleic acid which
XX hybridises to (I), and for determining the phenotype of cells in a sample
XX of cells from a patient. (I) is useful for determining the presence of
XX colon cancer in a cell or tissue type, for determining the presence or
XX state of other type of cancer, in antisense therapy, to generate
XX macroarrays on a solid surface, to identify a chromosome on which the
XX corresponding gene resides, and in tissue profiling, forensics, genetic
XX analysis, mapping and diagnostic applications. (I) can be used to raise
XX antibodies, and to screen for peptide analogues and antagonists.
XX
XX Sequence 571 BP; 136 A; 155 C; 131 G; 145 T; 4 other;
XX
Alignment Scores:
Pred. No.: 8.61e-29 Length: 571
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
09-833799-13B (1-57) x ABQ58818 (1-571)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 210 GCGCAAGAGCCAGTCAAGGTCCTCCACTAAGCTGCTGCTGCCCATTTATCTTG 269
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 270 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGACTGCCACGA 329
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 330 ATCAAGAAAGTGTGTGAAGGCTCTTGCGGGATGGCTGTTTCGTTCCTCCAG 380
RESULT 6
ABK83819
ID ABK83819 standard; cDNA; 2309 BP.
XX
XX ABK83819;
XX
XX 14-AUG-2002 (first entry)
XX
XX Human cDNA differentially expressed in granulocytic cells #390.
XX
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
XX fungal infection; sterile inflammatory disease; psoriasis;
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX adult respiratory distress syndrome; inflammatory bowel disease;
XX

```

KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX Homo sapiens.

XX WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US30821.

XX 03-OCT-2000; 2000US-237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -

XX Claim 1; SEQ ID No 390; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Alignment Scores:

Pred. No.:	4.77e-28	Length:	2309
Score:	327.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

09-833799-13B (1-57) x ABK83819 (1-2309)
 QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleu 20
 Db 1554 GCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCCATATCTTG 1613
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 1614 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACCTGACTGCCAGGA 1673
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 1674 ATCAAGAAAGTCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 1724

RESULT 7

ABL65816

ID ABL65816 standard; DNA; 2309 BP.

XX ABL65816;

AC ABL65816;

XX 15-MAY-2002 (first entry)

XX Lung cancer related gene sequence SEQ ID NO:4153.

DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 05-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-233133P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234009P.

XX 20-SEP-2000; 2000US-234034P.

XX 22-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 25-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

XX 25-SEP-2000; 2000US-235082P.

XX 25-SEP-2000; 2000US-235134P.

XX 26-SEP-2000; 2000US-235280P.

XX 26-SEP-2000; 2000US-235637P.

XX 26-SEP-2000; 2000US-235638P.

XX 27-SEP-2000; 2000US-235711P.

XX 27-SEP-2000; 2000US-235720P.

XX 27-SEP-2000; 2000US-235840P.

XX 27-SEP-2000; 2000US-235863P.

XX 28-SEP-2000; 2000US-236028P.

XX 28-SEP-2000; 2000US-236032P.

XX 28-SEP-2000; 2000US-236033P.

XX 28-SEP-2000; 2000US-236034P.

XX 28-SEP-2000; 2000US-236109P.

XX 29-SEP-2000; 2000US-236111P.

XX 29-SEP-2000; 2000US-236842P.

XX 29-SEP-2000; 2000US-236891P.

XX 02-OCT-2000; 2000US-237172P.

XX 02-OCT-2000; 2000US-237173P.

XX 02-OCT-2000; 2000US-237278P.

XX 02-OCT-2000; 2000US-237294P.

XX 02-OCT-2000; 2000US-237295P.

XX 02-OCT-2000; 2000US-237316P.

XX 03-OCT-2000; 2000US-237425P.

PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 4153; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Alignment Scores:
 Pred. No.: 4.77e-28 Length: 2309
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL65816 (1-2309)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleIleu 20
 |||||
 Db 1554 GGCACAGAGCCAGTCAAGGTCACGTCTCCACTAAGCCCTGCTCTGCCCCATTATCTTG 1613
 |||||

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 1614 ATCCGGTGGCCATGTTGAATCCGCCCTAACCGCTGCTTGAAGATACTAGCTGCCACGA 1673
 |||||

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 1674 ATCAAGAGTGTGTGAAGGCTGTTCGGGATGCGCTGTTTCGTTCCTCCAG 1724
 |||||

RESULT 8
 ABL66488
 ID ABL66488 standard; DNA; 2309 BP.
 XX
 XX ABL66488;
 AC
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4825.
 XX
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 XX gene; ds.
 OS Homo sapiens.
 XX
 XX WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 XX 30-MAY-2001; 2001WO-US10838.
 PF
 XX
 XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 XX (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 4825; 44pp; English.
 PS
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX

CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. MI can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of MI, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. MI can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Alignment Scores:
 Pred. No.: 4, 77e-28 Length: 2309
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL66488 (1-2309)

QY 1 AlaGlnGluProVallyysGlyProValSerThrLysProGlySerCysProIleLeu 20
 DB 1554 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCGCTGGCTCTGCCCATTTATCTTG 1613
 QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 1614 ATCCGGTGGCCGATGTGAATCCCGCTAACCGCTGCTTGAAGATACACTGACTGCCACGGA 1673
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 1674 ATCAAGAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTCGTTCCGCCAG 1724

RESULT 9
 AAQ56674
 ID AAQ56674 standard; DNA; 177 BP.
 AC AAQ56674;
 XX
 XX
 DT 23-SEP-1994 (first entry)
 DE Elafin derivative Leu25 DNA.

XX Elafin; derivative; mutation; transformation; E. coli; yeast;
 KW Bacillus subtilis; elastase inhibitor; oxidation; ss.
 XX Synthetic.

XX WO9404697-A.
 XX PD 03-MAR-1994.
 XX

PF 11-AUG-1993; 93WO-JP01133.

PR 11-AUG-1992; 92JP-0234085.

XX (TSUR) TSUMURA & CO.

PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
 XX

DR WPI; 1994-083211/10.
 DR P-PSDB; AAR48552.

XX Recombinant modified elafin with improved oxidation stability -
 PT has pharmaceutical use as an elastase inhibitor

PS Claim 1; Page 23; 35pp; Japanese.

XX The sequences given in AAQ56674-76 encode elafin derivatives. The DNA

CC encoding wild type elafin is mutated, inserted into a suitable
 CC vector and then used to transform E. coli, yeast, Bacillus subtilis
 CC or animal cells. The modified elafin is expressed when the
 CC transformed cells are cultured. The modified elafin are drugs with
 CC elastase inhibitor activity. They have improved oxidation stability
 CC over natural elafin and thus retain activity better under oxidative
 CC conditions.

SQ Sequence 177 BP; 39 A; 49 C; 48 G; 41 T; 0 other;

Alignment Scores:
 Pred. No.: 4, 55e-29 Length: 177
 Score: 324.00 Matches: 56
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 99.08% Indels: 0
 DB: 15 Gaps: 0

09-833799-13B (1-57) x AAQ56674 (1-177)

QY 1 AlaGlnGluProVallyysGlyProValSerThrLysProGlySerCysProIleLeu 20
 DB 1 GCACAGGAGACAGTTAAAGTCCGGTGTCCACCAACCGGGCTCTTGGCCGATTATCCTG 60
 QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 61 ATCCGGTGGCTTGTGCTGAACCGCGCGAACCGTGTCTGAAGACACTGACTGCCCGGGT 120
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 121 ATCAAAAATGCTGCGAAGGTCTTGGGTATGCGCATGCTTCGTCGCCGAG 171

RESULT 10

AAV28852

ID AAV28852 standard; cDNA to mRNA; 8598 BP.

AC AAV28852;

XX 04-AUG-1998 (first entry)

DE pPIC9/ELF25L cDNA construct containing the elafin 25L gene.

KW Elafin; elafin 25L gene; Pichia pastoris; expression vector;
 KW alcohol oxidase 1; controlling region; promoter; ds.

XX Synthetic.
 XX Pichia pastoris.

XX Key Location/Qualifiers

FT CDS 949..1377

FT /*tag= a

FT sig_peptide 949..1203

FT /*tag= b

FT mat_peptide 1204..1374

FT /*tag= c

XX JP10127292-A.

XX 19-MAY-1998.

XX 31-OCT-1996; 96JP-0304233.

XX 31-OCT-1996; 96JP-0304233.

XX (TSUR) TSUMURA & CO.

XX WPI; 1998-340667/30.

XX P-PSDB; AAW57237.

XX New elafin expression vector - used to transform microbial host(s)

XX for production of elafin commercially

XX Example 1; Page 11-15; 21pp; Japanese.

XX The present sequence represents a pPIC9/ELP25L cDNA construct containing
 CC the elafin 25L gene, used in an example of the present invention. The
 CC present invention describes: (1) an elafin (EL) expression vector
 CC comprising a gene encoding EL and a gene encoding alcohol oxidase 1
 CC controlling region (AOCR) which controls the expression of EL encoding
 CC gene; (2) a microbial host transformed with the vector; (3) DNA fragment
 CC for recombinant transformation comprising: (a) a first region homologous
 CC to the genomic DNA of the host; (b) a promoter region of (AOCR) gene; (c)
 CC a gene encoding signal peptide; (d) a linker designed to secrete a
 CC matured elafin; (e) a gene encoding EL; (f) selection marker gene; (g)
 CC stop codon, and (h) a second region homologous to a part of the genomic
 CC DNA of the host, and (4) a microbial host transformed with the DNA
 CC fragment of (3). The microbial hosts can be used to prepare EL
 CC commercially. The production of EL by the microbial hosts can be carried
 CC on a large scale.

SQ Sequence 8598 BP; 2306 A; 2025 C; 1920 G; 2347 T; 0 other;

Alignment Scores:
 Pred. No.: 5,31e-27 Length: 8598
 Score: 324.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 99.08% Indels: 0
 DB: 19 Gaps: 0

09-833799-13B (1-57) x AAV28852 (1-8598)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 DB 1204 GCTCAAGAACCCAGTTAAGGTCGGGTGTCGACCAACCGGGCTCTTGGCCGATTATCCTG 1263
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 1264 ATCCGCTGCGCTTTCGCTGAACCCGCGGAACCGTGTCTGTAAGACACTGACTGCCCGGT 1323
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 1324 ATCAAAAAATGCTGCGAAGGTTCTTGGCGTATGGCATGCTTCGTTCCGCAG 1374

RESULT 11
 AAQ56675
 ID AAQ56675 standard; DNA; 177 BP.
 XX
 AC AAQ56675;
 XX
 DT 23-SEP-1994 (first entry)
 XX
 DE Elafin derivative Val25 DNA.
 XX
 KW Elafin; derivative; mutation; transformation; E. coli; yeast;
 KW Bacillus subtilis; elastase inhibitor; oxidation; ss.
 XX
 OS Synthetic.
 XX
 PN WO9404697-A.
 XX
 PD 03-MAR-1994.
 XX
 PF 11-AUG-1993; 93WO-JP01133.
 XX
 PR 11-AUG-1992; 92JP-0234085.
 XX
 PA (TSUR) TSUMURA & CO.

PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
 XX WPI; 1994-083211/10.
 DR P-PSDB; AAR48553.
 XX
 PT Recombinant modified elafin with improved oxidation stability -
 PT has pharmaceutical use as an elastase inhibitor

XX Claim 1; Page 24; 35pp; Japanese.
 PS
 CC The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
 CC encoding wild type elafin is mutated, inserted into a suitable
 CC vector and then used to transform E. coli, yeast, Bacillus subtilis
 CC or animal cells. The modified elafin is expressed when the
 CC transformed cells are cultured. The modified elafin are drugs with
 CC elastase inhibitor activity. They have improved oxidation stability
 CC over natural elafin and thus retain activity better under oxidative
 CC conditions.

SQ Sequence 177 BP; 39 A; 50 C; 48 G; 40 T; 0 other;

Alignment Scores:
 Pred. No.: 5,94e-29 Length: 177
 Score: 323.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 99.78% Indels: 0
 DB: 15 Gaps: 0

09-833799-13B (1-57) x AAQ56675 (1-177)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 DB 1 GCACAGGAACCCAGTTAAAGTCCGGTGTGCGACCAACCGGGCTCTTGGCCGATTATCCTG 60
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 61 ATCCGCTGCGCTTCTGTAACCCGCGGAACCGTGTCTGTAAGACACTGACTGCCCGGT 120
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 121 ATCAAAAAATGCTGCGAAGGTTCTTGGCGTATGGCATGCTTCGTTCCGCAG 171

RESULT 12
 AAQ56676
 ID AAQ56676 standard; DNA; 177 BP.
 XX
 AC AAQ56676;
 XX
 DT 23-SEP-1994 (first entry)
 XX
 DE Elafin derivative Ile25 DNA.
 XX
 KW Elafin; derivative; mutation; transformation; E. coli; yeast;
 KW Bacillus subtilis; elastase inhibitor; oxidation; ss.
 XX
 OS Synthetic.
 XX
 PN WO9404697-A.
 XX
 PD 03-MAR-1994.
 XX
 PF 11-AUG-1993; 93WO-JP01133.
 XX
 PR 11-AUG-1992; 92JP-0234085.
 XX
 PA (TSUR) TSUMURA & CO.

PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
 XX WPI; 1994-083211/10.
 DR P-PSDB; AAR48554.
 XX
 PT Recombinant modified elafin with improved oxidation stability -
 PT has pharmaceutical use as an elastase inhibitor

PS Claim 1; Page 25; 35pp; Japanese.

CC The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
 CC encoding wild type elafin is mutated, inserted into a suitable

CC vector and then used to transform E. coli, yeast, Bacillus subtilis
 CC or animal cells. The modified elafin is expressed when the
 CC transformed cells are cultured. The modified elafin are drugs with
 CC elastase inhibitor activity. They have improved oxidation stability
 CC over natural elafin and thus retain activity better under oxidative
 CC conditions.

XX
 SQ Sequence 177 BP; 40 A; 49 C; 47 G; 41 T; 0 other;

Alignment Scores:
 Pred. No.: 5,948-29 Length: 177
 Score: 323.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 98.78% Indels: 0
 DB: 15 Gaps: 0

09-833799-13B (1-57) x AAQ56676 (1-177)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 DB 1 GCACAGGAACCAAGTTAAAGTCCGGTGTGACCAACCGGGCTCTTGCCGATTATCTCG 60
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 61 ATCCGCTGGCTATTCTGAACCCGCGAACCGTGTCTGAAAGACACTGACTGCCCGGT 120
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 121 ATCAAAATATGTCGAAAGTTCTTGGGTATGGCATGCTTCTGTTCCGAG 171

RESULT 13

AAQ28757
 ID AAQ28757 standard; DNA; 321 BP.

AC AAQ28757;

DT 25-FEB-1993 (first entry)

DE Partial sequence of tumour suppressor gene U9.

KW Can19; tumour suppressor gene; cancer; therapy; ss.

OS Homo sapiens.

PN W09215602-A.

PD 17-SEP-1992.

PF 28-FEB-1992; 92WO-US01624.

PR 28-FEB-1991; 91US-0662216.

PA (DAND) DANA FABER CANCER INST INC.

PI Sager R;

DR WPI; 1992-331663/40.

PT Diagnoses and treatment of cancer - using candidate tumor suppressor
 PT genes or the corresp. antibodies.

PS Claim 29; Page 37-38; 54pp; English.

CC An adaptation of the subtractive hybridization technique was used
 CC which utilizes a biotinylation-based subtraction procedure instead
 CC of hydroxyapatite as previously used. In this procedure, a single
 CC strand phagemid cDNA library from normal cell polyA+ mRNA is
 CC hybridized with excess biotinylated tumor polyA+ mRNA, and the
 CC resulting double stranded sequences are removed by binding to
 CC streptavidin. The remaining single-stranded phagemid cDNAs are
 CC converted to double-stranded form and used to transform bacterial
 CC host cells. The resulting subtracted cDNA library is differentially

CC screened with total cDNA from normal and tumor cells. This method
 CC produced some 20 additional cloned cDNAs. Also found by this
 CC method were several genes which, on the basis of the partial DNA
 CC sequences appear to be novel sequences not previously entered
 CC into GENBANK. The portion of the cDNAs so sequenced represents
 CC part of the coding region and/or part of the 3' untranslated region
 CC of each cDNA (see Q28749-58).

XX
 SQ Sequence 321 BP; 81 A; 80 C; 84 G; 76 T; 0 other;

Alignment Scores:
 Pred. No.: 1,618-28 Length: 321
 Score: 322.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.47% Indels: 0
 DB: 13 Gaps: 0

09-833799-13B (1-57) x AAQ28757 (1-321)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 DB 144 GCACAGGAACCAAGTCAAGGTCAGTCCACTTAAGCTGGCTCTGCCCATATCTTG 203
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 DB 204 ATCCGCTGGCTATTCTGAACCCGCGAACCGTGTCTGAAAGACACTGACTGCCCGGA 263
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 DB 264 ATCAAGAAGTGTGTGAAGGCTCTGCGGGATGGCTGTTTCGTTCC 311

RESULT 14

AAQ44862
 ID AAQ44862 standard; cDNA to mRNA; 737 BP.

AC AAQ44862;

DT 17-OCT-1994 (first entry)

DE SPAl gene.

KW SPAl; sodium ion, potassium ion, ATP-ase inhibitor peptide; cardiac;
 KW vascular disease; PCR; polymerase chain reaction; ss.

OS Sus scrofa.

XX
 FH Key Location/Qualifiers

FT CDS 1..570

FT /*tag= a

FT /product= SPAl

XX JP06049098-A.

XX 22-FEB-1994.

XX 29-JUL-1992; 92JP-0202286.

XX 29-JUL-1992; 92JP-0202286.

XX (EISA) EISAI CO LTD.

XX WPI; 1994-097819/12.

XX P-PSDB; AAR50334.

XX Sodium ion, potassium ion ATP-ase inhibitor peptide and corresp.
 XX DNA - useful to treat cardiac and vascular disease

XX Claim 1; Page 6-7; 8pp; Japanese.

XX The sequence (AAQ44862) encodes the SPAl peptide which is useful
 XX as a drug for treating cardiac and vascular diseases. The peptide
 XX is found as a pre-sequence (AAR50335) which is amplified using

CC primers (AAQ44863-67)
 XX Sequence 737 BP; 205 A; 192 C; 182 G; 158 T; 0 other;
 SQ Alignment Scores:
 Pred. No.: 1.26e-11 Length: 737
 Score: 179.50 Matches: 31
 Percent Similarity: 61.90% Conservativity: 8
 Best Local Similarity: 49.21% Mismatches: 17
 Query Match: 54.89% Indels: 7
 DB: 15 Gaps: 1

09-833799-13B (1-57) x AAQ44862 (1-737)
 QY 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
 DB 379 CAAGATCCAGTCAAGCCCACTCCAGTCCAGTCCAGTCTCTCTCTTAAGCGTGGC 438
 QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
 DB 439 CACTGCCCTAGGATCTTTTTCGTTGCCCGCTGAGCAATCCCTCTAACAAAGTTGGAGA 498
 QY 35 AspThrAspCysProGlyLleLysLysCysGluGlySerCysGlyMetAlaCysPhe 54
 DB 499 GATTATGACTGTCAGGGGTCAAGAAAGTCTGTGAAGGCTTTTGGCGGAGGATTGTTG 558
 QY 55 ValProGln 57
 DB 559 TATCCCAAG 567

RESULT 15
 ABQ73674
 ID ABQ73674 standard; DNA; 396 BP.
 AC ABQ73674;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Murine SLPI sense riboprobe SEQ ID NO:20.
 KW Murine; serum amyloid; SAA3; riboprobe; LST-1; follistatin-like 1;
 KW FST1; SLPI; secretory leukocyte protease inhibitor; legumain; SOCS3;
 KW rheumatoid arthritis; CISH3; RAGE; AGER; LY17; hydroxymethyl glutarate;
 KW HMG-1; glucocorticoid leucine zipper; GILZ; PTPN-18; GADD-45A; GADD-45B;
 KW PRSC1; lipocalin 2; Lcn2; glucose phosphate isomerase; GPI; Spil;
 KW serine protease inhibitor; TSG-6; ss.
 XX
 OS Mus sp.
 XX
 PN WO200248310-A2.
 XX
 PD 20-JUN-2002.
 XX
 PF 17-DEC-2001; 2001WO-US48968.
 XX
 PR 15-DEC-2000; 2000US-255861P.
 XX
 PA (GENE-) GENETICS INST LLC.
 XX
 PI Pittman DD, Feldman JL, Shields KM, Trepicchio WL;
 XX
 DR WPI; 2002-583494/62.
 XX

Determining difference between expression levels of genes
 PT characteristic of rheumatoid arthritis in cell and reference level, by
 PT comparing expression levels of the genes determined in a cell with
 PT reference level -
 XX
 PS Example 3; Page 88; 296pp; English.
 XX
 CC The present invention describes a method (M1) for determining the
 CC difference between levels of expression of a number of genes
 CC characteristic of rheumatoid arthritis (RA) in cell and reference levels

of expression of the genes, comprising: (a) providing RNA from a cell;
 (b) determining levels of RNA of a number of genes (G) characteristic
 of RA including a number of genes selected from SOCS3 (CISH3), RAGE
 (AGER), LST-1 (LY17), serum amyloid (SAA) 1-3, hydroxymethyl glutarate
 (HMG)-1, S100 A8, A9, and A12, secretory leukocyte protease inhibitor
 (SLPI), glucocorticoid leucine zipper (GILZ), PTPN-18, GADD-45A and B,
 legumain (PRSC1), follistatin-like 1 (FST1), lipocalin 2 (Lcn2), glucose
 phosphate isomerase (GPI), serine protease inhibitor (Spil), and TSG-6
 to obtain the levels of expression of the genes in the cell; and
 (c) comparing the levels of expression of the genes in the cell to a
 set of reference levels of expression of the genes, to determine the
 difference between levels of expression of the number of genes
 characteristic of RA in the cell and reference levels of expression of
 the genes. M1 is useful for determining whether a subject has or is
 likely to develop RA, or for determining whether a therapy for RA is
 effective in a subject having RA who is receiving the therapy. M1 is
 also useful for determining the stage of RA in a subject, for
 determining the efficacy of a therapy in a subject having RA, and for
 determining the likelihood of success of a particular therapy in subject
 having RA. The present sequence represents a murine SLPI sense riboprobe,
 CC which is used in an example from the present invention.
 XX
 SQ Sequence 396 BP; 94 A; 98 C; 118 G; 86 T; 0 other;

Alignment Scores:
 Pred. No.: 8.03e-10 Length: 396
 Score: 161.00 Matches: 28
 Percent Similarity: 58.49% Conservativity: 3
 Best Local Similarity: 52.83% Mismatches: 22
 Query Match: 49.24% Indels: 0
 DB: 24 Gaps: 0

09-833799-13B (1-57) x ABQ73674 (1-396)
 QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
 DB 232 CCCATTCGCAACACAGTGTGGAGGAGCCCTGGAGGTGGCTCAAAACTCAGCAAGATGT 291
 QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLleLysLys 43
 DB 292 ATGATGCTTTAAACCTCCCAATGCTGCCAGAGGAGCGGCGAGTGTGACGCAATACAG 351
 QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 DB 352 TGCTGTGAGGTATATGTGGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390

RESULT 16
 ABQ73677/c
 ID ABQ73677 standard; DNA; 396 BP.
 XX
 AC ABQ73677;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Murine SLPI antisense riboprobe SEQ ID NO:23.
 XX
 KW Murine; serum amyloid; SAA3; riboprobe; LST-1; follistatin-like 1;
 KW FST1; SLPI; secretory leukocyte protease inhibitor; legumain; SOCS3;
 KW rheumatoid arthritis; CISH3; RAGE; AGER; LY17; hydroxymethyl glutarate;
 KW HMG-1; glucocorticoid leucine zipper; GILZ; PTPN-18; GADD-45A; GADD-45B;
 KW PRSC1; lipocalin 2; Lcn2; glucose phosphate isomerase; GPI; Spil;
 KW serine protease inhibitor; TSG-6; ss.
 XX
 OS Mus sp.
 XX
 PN WO200248310-A2.
 XX
 PD 20-JUN-2002.
 XX
 PF 17-DEC-2001; 2001WO-US48968.
 XX
 PR 15-DEC-2000; 2000US-255861P.
 XX


```

XX DE Mouse IMC carcinoma cell IMC-HA1 clone 28-1#3 CDNA.
XX AC
XX KW Mouse; carcinoma cell; IMC-HA1; cancer; metastasis; CMAP; inhibitor;
XX KW Cancer metastasis associated protein; ss.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
XX FT CDS 37..432
XX FT /*tag= a
XX PN WO9845431-A1.
XX PD 15-OCT-1998.
XX PF 07-APR-1998; 98WO-JP01592.
XX PR 08-APR-1997; 97JP-0105333.
XX PA (BANY ) BANYU PHARM CO LTD.
XX PI Arakawa H, Morita M, Ohta M;
XX WPI; 1999-080732/07.
XX DR P-PSDB; AAW98908.
XX PT Protein associated with cancer metastasis and gene encoding it -
XX PT useful for screening for potential inhibitors of cancer metastasis
XX PS Claim 2; Page 41-42; 74pp; Japanese.
XX CC The present invention provides gene sequences associated with cancer
XX CC metastasis which are isolated from mouse IMC carcinoma cells by
XX CC detection of their higher expression in IMC-HM cell lines than in
XX CC IMC-LM cell lines using differential display of the mRNA in these cells.
XX CC The gene sequences can be used for the screening of potential inhibitors
XX CC of cancer metastasis by either: bringing into contact with the cancer
XX CC metastasis associated protein (CMAP) and determining the degree of
XX CC binding; or creating a transformant cell line which expresses CMAP and
XX CC measuring the degree of expression of CMAP using an antibody recognising
XX CC the protein, either in the presence or absence of the potential
XX CC inhibitor. IMC-HM cells transformed with antisense CMAP DNA show a
XX CC lowered ability to metastasise. The present sequence represents a
XX CC specifically claimed gene sequence from the present invention.
XX SQ Sequence 691 BP; 160 A; 186 C; 183 G; 162 T; 0 other;

Alignment Scores:
Pred. No.: 1.59e-09 Length: 691
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 20 Gaps: 0

09-833799-13B (1-57) x AAX18514 (1-691)
Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
Db 268 CCCATTCGCAACACAGTGTGGAGGAGCCTGGGAGGTGCGTCAAACTCAGGCAAGATGT 327
Qy 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 328 ATGATGCTTAACTCCCAATGTCTGCAGAGGACGGCGAGGTGACGCGCAATACAG 387
Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 388 TGCTGTGAGGGTATATGTGGAAAGTCTGCCTGCCCCCG 426

RESULT 19
AAX18515
ID AAX18515 standard; cDNA; 1114 BP.

```

```

XX AAX18515;
XX 05-MAY-1999 (first entry)
XX Mouse IMC carcinoma cell IMC-HA1 clone #8.323 cDNA.
XX KW Mouse; carcinoma cell; IMC-HA1; cancer; metastasis; CMAP; inhibitor;
XX KW Cancer metastasis associated protein; ss.
XX OS Mus musculus.
XX PN WO9845431-A1.
XX PD 15-OCT-1998.
XX PF 07-APR-1998; 98WO-JP01592.
XX PR 08-APR-1997; 97JP-0105333.
XX PA (BANY ) BANYU PHARM CO LTD.
XX PI Arakawa H, Morita M, Ohta M;
XX WPI; 1999-080732/07.
XX PT Protein associated with cancer metastasis and gene encoding it -
XX PT useful for screening for potential inhibitors of cancer metastasis
XX PS Claim 2; Page 44-45; 74pp; Japanese.
XX CC The present invention provides gene sequences associated with cancer
XX CC metastasis which are isolated from mouse IMC carcinoma cells by
XX CC detection of their higher expression in IMC-HM cell lines than in
XX CC IMC-LM cell lines using differential display of the mRNA in these cells.
XX CC The gene sequences can be used for the screening of potential inhibitors
XX CC of cancer metastasis by either: bringing into contact with the cancer
XX CC metastasis associated protein (CMAP) and determining the degree of
XX CC binding; or creating a transformant cell line which expresses CMAP and
XX CC measuring the degree of expression of CMAP using an antibody recognising
XX CC the protein, either in the presence or absence of the potential
XX CC inhibitor. IMC-HM cells transformed with antisense CMAP DNA show a
XX CC lowered ability to metastasise. The present sequence represents a
XX CC specifically claimed gene sequence from the present invention.
XX SQ Sequence 1114 BP; 271 A; 299 C; 290 G; 254 T; 0 other;

Alignment Scores:
Pred. No.: 2.85e-09 Length: 1114
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 20 Gaps: 0

09-833799-13B (1-57) x AAX18515 (1-1114)
Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
Db 685 CCCATTCGCAACACAGTGTGGAGGAGCCTGGGAGGTGCGTCAAACTCAGGCAAGATGT 744
Qy 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 745 ATGATGCTTAACTCCCAATGTCTGCAGAGGACGGCGAGGTGACGCGCAATACAG 804
Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 805 TGCTGTGAGGGTATATGTGGAAAGTCTGCCTGCCCCCG 843

RESULT 20
AAC97526
ID AAC97526 standard; DNA; 180 BP.

```


SQ Sequence 183 BP; 43 A; 45 C; 51 G; 44 T; 0 other;

Alignment Scores: 1.68e-08 Length: 183
 Pred. No.: 146.00 Matches: 27
 Score: 55.36% Conservative: 4
 Percent Similarity: 48.21% Mismatches: 23
 Best Local Similarity: 44.65% Indels: 2
 Query Match: 20 Gaps: 1
 DB:

09-833799-13B (1-57) x AAX16273 (1-183)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 4 GATCCTGTTGACACCCCAACACCAAGAGGAGGAGCTGGGAAGTCCCGAGTACTTAT 63
 Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 64 GGCCAATGTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCACGCGT 123
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 124 GACTTGAAGTCTGATGGCGATGTGTGGAAATCCTCGCTTCCCT 171

RESULT 22

AAI67584
 ID AAI67584 standard; cDNA; 183 BP.

XX AC AAI67584;

XX DT 11-FEB-2002 (first entry)

XX DE Secretory leukocyte protease inhibitor (SLPI) analogue cDNA.

XX KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
 XX KW trypsin; secretory leukocyte protease inhibitor; SLPI; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX CDS 1..183
 FT /*tag= a
 FT /transl_except= "(pos: 22..24, aa: Asn)"

XX PN US6291662-B1.

XX PD 18-SEP-2001.

XX PF 22-SEP-1998; 98US-0158085.

XX PR 30-MAR-1987; 87US-0031846.

XX PR 06-AUG-1990; 90US-0563832.

XX PR 22-JUL-1994; 94US-0279056.

XX PR 05-DEC-1984; 84US-0678222.

XX PR 02-DEC-1985; 85US-0803471.

XX PR 29-JUL-1986; 86US-0890526.

XX PA (AMGE-) AMGEN INC.

XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX DR WPI; 2001-637974/73.

XX DR P-PSDB; AAG65997.

XX FT New DNA sequences, useful in recombinant DNA techniques for directing
 the production of a serine protease inhibitor protein, e.g. leukocyte
 elastase or trypsin

XX PS Example 9; Column 42; 37pp; English.

XX CC The invention relates to a DNA sequence encoding an analog of a mammalian
 CC serine protease inhibitor protein. The analog comprises at least eight
 CC cysteine residues and possesses serine protease inhibitor activity, and

CC at least one operational element consisting of a promoter, an operator,
 CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
 CC or a terminator codon. The DNA is useful in recombinant DNA techniques
 CC for directing the production of a serine protease inhibitor protein, the
 CC e.g. leukocyte elastase or trypsin. The present sequence represents the
 CC cDNA encoding a secretory leukocyte protease inhibitor (SLPI) analogue
 CC that has chymotrypsin and elastase inhibitory activity.

XX SQ Sequence 183 BP; 43 A; 44 C; 51 G; 45 T; 0 other;

Alignment Scores: 1.68e-08 Length: 183
 Pred. No.: 146.00 Matches: 27
 Score: 55.36% Conservative: 4
 Percent Similarity: 48.21% Mismatches: 23
 Best Local Similarity: 44.65% Indels: 2
 Query Match: 22 Gaps: 1
 DB:

09-833799-13B (1-57) x AAI67584 (1-183)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 4 GATCCTGTTGACACCCCAACACCAAGAGGAGGAGCTGGGAAGTCCCGAGTACTTAT 63

Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 64 GGCCAATGTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCACGCGT 123

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 124 GACTTGAAGTCTGATGGCGATGTGTGGAAATCCTCGCTTCCCT 171

RESULT 23

AAAN90354

ID AAAN90354 standard; DNA; 194 BP.

XX AC AAAN90354;

XX DT 01-NOV-1989 (first entry)

XX DE Sequence encoding fragment of human polymorphonuclear leukocyte
 XX DE elastase inhibiting protein.

XX KW Elastase inhibitor; human polymorphonuclear leukocyte elastase inhibiting
 XX KW protein.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers
 XX FT misc_feature 1..5
 FT /*tag= a
 FT misc_feature 7..12
 FT /*tag= b
 FT misc_feature 25..30
 FT /*tag= c
 FT misc_feature 56..61
 FT /*tag= d
 FT misc_feature 120..125
 FT /*tag= e
 FT misc_feature 185..189
 FT /*tag= f
 FT misc_feature 190..194
 FT /*tag= g

XX PN WO8906239-A.

XX PD 13-JUL-1989.

XX PF 28-DEC-1988; 88WO-JP01342.

XX PR 28-DEC-1987; 87JP-0330219.

XX PA (TEIJ) TEIJIN LTD.

XX 22-JUL-1994; 94US-0279056.
 XX 30-MAR-1987; 87US-0031846.
 PR 05-DEC-1984; 84US-0678822.
 PR 29-JUL-1986; 86US-0890526.
 PR 03-SEP-1986; 86US-0903471.
 PR 06-AUG-1990; 90US-0563832.
 PR 22-JUL-1994; 94US-0279056.
 XX (AMGE-) AMGEN INC.
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX WPI; 1999-166640/14.
 XX New DNA sequence encoding mammalian serine protease inhibitor -
 PT useful for recombinantly producing inhibitors with different
 PT specificities and treating diseases such as emphysema, arthritis,
 PT muscular dystrophy, and tumour invasion
 XX Example 1; Column 23; 37pp; English.
 XX The present invention describes a DNA sequence (A) encoding an analogue
 CC of a mammalian serine protease inhibitor (B). The DNA sequences and
 CC recombinant methods allow manufacture of a class of inhibitors of e.g.
 CC cathepsin G, elastase, and trypsin, with different specificities. The
 CC recombinant serine protease product can be directed to act
 CC intracellularly or extracellularly and is useful in treating conditions
 CC caused by a disturbance in the native protease/protease inhibitor
 CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
 CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
 CC allows (B) to be recombinantly produced in sufficient quantities and
 CC purities so as to provide economical pharmaceutical compositions. (B) is
 CC resistant to heat, acid, and a variety of proteolytic enzymes, is
 CC thermodynamically stable in extracellular conditions, and exhibits a
 CC high degree of self assembly forming an active tertiary structure in the
 CC absence of biochemical stimuli. The present sequence represents a DNA
 CC sequence which is used to direct manufacture of recombinant serine
 CC protease inhibitors.
 XX SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.52e-08 Length: 324
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 20 Gaps: 1
 09-833799-13B (1-57) x AAX16194 (1-324)
 Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
 Db 145 GATCGGTTGATACCCGACCCGACTCTGCGAAACCGGGTAAATGCCCGTAACCTAT 204
 Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 205 GCCCAGTGTCTGATGCTGAACCCCGCAACTCTTCGCGAAATGGACGCCAGTGTAAACGA 264
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 265 GATCTGAATGCTGATGGGTATGTGGCGCAATCTTGTTTCCCG 312
 RESULT 28
 ID AAX16236
 ID AAX16236 standard; DNA; 324 BP.
 XX AAX16236;
 AC AAX16236;
 XX 19-APR-1999 (first entry)
 DT
 XX

DE Serine protease inhibitor direct manufacturing DNA sequence (SLPT).
 XX Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
 KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
 KW acute leukemia; ss.
 XX Homo sapiens.
 OS Synthetic.
 XX US5871956-A.
 XX 16-FEB-1999.
 XX 22-JUL-1994; 94US-0279056.
 XX 30-MAR-1987; 87US-0031846.
 PR 05-DEC-1984; 84US-0678822.
 PR 29-JUL-1986; 86US-0890526.
 PR 03-SEP-1986; 86US-0903471.
 PR 06-AUG-1990; 90US-0563832.
 PR 22-JUL-1994; 94US-0279056.
 XX (AMGE-) AMGEN INC.
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX WPI; 1999-166640/14.
 XX New DNA sequence encoding mammalian serine protease inhibitor -
 PT useful for recombinantly producing inhibitors with different
 PT specificities and treating diseases such as emphysema, arthritis,
 PT muscular dystrophy, and tumour invasion
 XX Example 2; Column 29; 37pp; English.
 XX The present invention describes a DNA sequence (A) encoding an analogue
 CC of a mammalian serine protease inhibitor (B). The DNA sequences and
 CC recombinant methods allow manufacture of a class of inhibitors of e.g.
 CC cathepsin G, elastase, and trypsin, with different specificities. The
 CC recombinant serine protease product can be directed to act
 CC intracellularly or extracellularly and is useful in treating conditions
 CC caused by a disturbance in the native protease/protease inhibitor
 CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
 CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
 CC allows (B) to be recombinantly produced in sufficient quantities and
 CC purities so as to provide economical pharmaceutical compositions. (B) is
 CC resistant to heat, acid, and a variety of proteolytic enzymes, is
 CC thermodynamically stable in extracellular conditions, and exhibits a
 CC high degree of self assembly forming an active tertiary structure in the
 CC absence of biochemical stimuli. The present sequence represents a DNA
 CC sequence which is used to direct manufacture of recombinant serine
 CC protease inhibitors.
 XX SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.52e-08 Length: 324
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 20 Gaps: 1
 09-833799-13B (1-57) x AAX16236 (1-324)
 Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
 Db 145 GATCGGTTGATACCCGACCCGACCCGACTCTGCGAAACCGGGTAAATGCCCGTAACCTAT 204
 Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 205 GCCCAGTGTCTGATGCTGAACCCCGCAACTCTTCGCGAAATGGACGCCAGTGTAAACGA 264

CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Alignment Scores: 7.52e-08 Length: 324
 Pred. No.: 143.00 Matches: 27
 Score: 55.36% Conservative: 4
 Percent Similarity: 48.21% Mismatches: 23
 Best Local Similarity: 43.73% Indels: 1
 Query Match: 21 Gaps: 2

09-833799-13B (1-57) x AAC97528 (1-324)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
 Db 145 GATCCGGTTGATACCCGGAACCCGACTCGTGAACACCGGTAATGCCCGTAACCTAT 204
 QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 205 GGCACAGTCTGATGCTGAACCCCGCAACTTCTCGAAATGACGCCAGTGTAAACGA 264
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 265 GATCTGAAATGCTGATGGTATGTGCGCAAAATCTTGTGTTCCCG 312

RESULT 31

AAC97579
 ID AAC97579 standard; DNA; 324 BP.

XX AAC97579;

XX 27-FEB-2001 (first entry)

XX DNA encoding a serine protease inhibitory peptide.

XX Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.

XX Synthetic.

XX US6132990-A.

XX 17-OCT-2000.

XX 07-JUN-1991; 91US-0712354.

XX 03-JAN-1989; 89US-0293042.

PR 06-DEC-1984; 84US-0678822.

PR 02-DEC-1985; 85US-0803471.

PR 29-JUL-1986; 86US-0890526.

PR 30-MAR-1987; 87US-0031846.

PR 04-AUG-1987; 87US-0082962.

XX (AMGE-) AMGEN BOULDER INC.

XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX WPI; 2000-678667/66.

XX New serine protease inhibitors and DNA sequences for treating a

PT protease-mediated condition or tissue destruction e.g. emphysema or

PT tumor invasion and for recombinant production of inhibitors

XX Example 2; Column 30-32; 47pp; English.

XX This invention relates to new purified and isolated mammalian serine

CC protease inhibitor proteins which comprise at least 8 cysteine residues

CC and no more than 107 amino acids. The protease inhibitors are capable of

CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Alignment Scores:

Pred. No.: 7.52e-08 Length: 324
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97579 (1-324)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20

Db 145 GATCCGGTTGATACCCGGAACCCGACTCGTGAACACCGGTAATGCCCGTAACCTAT 204

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 205 GGCACAGTCTGATGCTGAACCCCGCAACTTCTCGAAATGACGCCAGTGTAAACGA 264

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 265 GATCTGAAATGCTGATGGTATGTGCGCAAAATCTTGTGTTCCCG 312

RESULT 32

AA167488

ID AA167488 standard; DNA; 324 BP.

XX AA167488;

XX 11-FEB-2002 (first entry)

XX Recombinant serine protease inhibitor DNA.

XX Serine protease inhibitor protein; recombinant; leukocyte elastase;
 KW trypsin; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..324

FT /*tag= a

FT /product= "serine protease inhibitor"

XX US6291662-B1.

XX 18-SEP-2001.

XX 22-SEP-1998; 98US-0158085.

XX 30-MAR-1987; 87US-0031846.

PR 06-AUG-1990; 90US-0563832.

PR 22-JUL-1994; 94US-0279056.

PR 05-DEC-1984; 84US-0678222.

PR 02-DEC-1985; 85US-0803471.

PR 29-JUL-1986; 86US-0890526.

PA (AMGE-) AMGEN INC.
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 PI
 XX WPI; 2001-637974/73.
 DR P-PSDB; AAG65993.
 XX
 XX New DNA sequences, useful in recombinant DNA techniques for directing
 PT the production of a serine protease inhibitor protein, e.g. leukocyte
 PT elastase or trypsin
 XX
 XX Example 1; Column 4; 37pp; English.
 XX
 CC The invention relates to a DNA sequence encoding an analog of a mammalian
 CC serine protease inhibitor protein. The analog comprises at least eight
 CC cysteine residues and possesses serine protease inhibitor activity, and
 CC at least one operational element consisting of a promoter, an operator,
 CC or a terminator codon. The DNA is useful in recombinant DNA techniques
 CC for directing the production of a serine protease inhibitor protein,
 CC e.g. leukocyte elastase or trypsin. The present sequence represents the
 CC coding strand of a DNA that directs the manufacture of a recombinant
 CC serine protease inhibitor.
 XX
 SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.52e-08 Length: 324
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 22 Gaps: 1
 09-833799-13B (1-57) x AAI67488 (1-324)
 QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCGGGTAATCCCGGTAACCTAT 204
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 205 GGCCAGTGTCTGATCGTGAACCCCGCAACTTCTCGAAATGGACGCCAGCTGTAAACGA 264
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 265 GATCTGAAATGCTGATCGGTATGTGCGGCAATCTTGTGTTCCCG 312
 RESULT 33
 AAI67489
 ID AAI67489 standard; DNA; 324 BP.
 XX
 AC AAI67489;
 XX
 DT 11-FEB-2002 (first entry)
 XX
 DE Recombinant secretory leukocyte protease inhibitor (SLPI) DNA.
 XX
 KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
 KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
 XX
 OS Homo sapiens.
 XX
 XX US6291662-B1.
 PN
 XX
 XX 18-SEP-2001.
 PD
 XX
 XX 22-SEP-1998; 98US-0158085.
 PF
 XX
 XX 30-MAR-1987; 87US-0031846.
 PR
 XX 06-AUG-1990; 90US-0563832.
 PR
 XX 22-JUL-1994; 94US-0273056.
 PR
 XX *05-DEC-1984; 84US-0678222.
 PR

PR 02-DEC-1985; 85US-0803471.
 PR 29-JUL-1986; 86US-0890526.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 PI
 XX WPI; 2001-637974/73.
 DR P-PSDB; AAG65996.
 XX
 XX New DNA sequences, useful in recombinant DNA techniques for directing
 PT the production of a serine protease inhibitor protein, e.g. leukocyte
 PT elastase or trypsin
 XX
 XX Disclosure; Column 5; 37pp; English.
 PS
 XX
 CC The invention relates to a DNA sequence encoding an analog of a mammalian
 CC serine protease inhibitor protein. The analog comprises at least eight
 CC cysteine residues and possesses serine protease inhibitor activity, and
 CC at least one operational element consisting of a promoter, an operator,
 CC or a terminator codon. The DNA is useful in recombinant DNA techniques
 CC for directing the production of a serine protease inhibitor protein,
 CC e.g. leukocyte elastase or trypsin. The present sequence represents the
 CC coding strand of a DNA that directs the manufacture of a recombinant
 CC secretory leukocyte protease inhibitor (SLPI).
 XX
 SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.52e-08 Length: 324
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 22 Gaps: 1
 09-833799-13B (1-57) x AAI67489 (1-324)
 QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCGGGTAATCCCGGTAACCTAT 204
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 205 GGCCAGTGTCTGATCGTGAACCCCGCAACTTCTCGAAATGGACGCCAGCTGTAAACGA 264
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 265 GATCTGAAATGCTGATCGGTATGTGCGGCAATCTTGTGTTCCCG 312
 RESULT 34
 ABA83108
 ID ABA83108 standard; DNA; 399 BP.
 XX
 AC ABA83108;
 XX
 DT 08-FEB-2002 (first entry)
 XX
 DE Secretory leukocyte protease inhibitor ovarian tumour marker gene, #53.
 XX
 KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 XX gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.

XX WO200175177-A2.
PN 11-OCT-2001.
XX
PD
XX
PF 03-APR-2001; 2001WO-US10947.
XX
PR 03-APR-2000; 2000US-194336P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
PI
XX WPI; 2001-626450/72.
DR P-PSDB; ABB50282.
XX
XX Detecting and identifying ovarian tumor, identifying increased risk for
PT developing ovarian cancer, and determining effectiveness of ovarian
PT cancer treatment, by measuring expression level of ovarian tumor marker
PT gene
XX
XX Claim 23; Page 98; 140pp; English.
XX
XX The invention relates to methods for diagnosing and prognosing ovarian
CC tumors in an individual via the detection and measurement of the
CC expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
CC ABA83181 and ABA83183). The methods of the invention are useful for
CC detecting an ovarian tumor in a patient, for identifying an individual
CC at increased risk for developing ovarian cancer, in prognostic tests for
CC assessing the relative severity of ovarian cancer, in tests for
CC monitoring a patient in remission from ovarian cancer, and in tests for
CC monitoring disease status in a patient being treated for ovarian cancer.
CC The methods can additionally be used to identify a particular tumor as
CC being an ovarian tumor (i.e., an epithelial ovarian tumor selected from
CC serous cystadenoma, borderline serous tumor, serous cystadenocarcinoma,
CC mucinous cystadenoma, borderline mucinous tumor, mucinous
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC tumor. The ovarian tumor marker genes of the invention were identified
CC using SAGE (serial analysis of gene expression) and were found to be
CC overexpressed in a broad variety of ovarian epithelial tumor cells
CC relative to normal ovarian epithelial cells. The marker genes are
CC implicated in immune response pathways, in the regulation of cell
CC proliferation and in protein folding, and many of these are membrane-
CC localised or secreted. In addition to their use as diagnostic and
CC prognostic markers, the ovarian tumor marker genes or their encoded
CC proteins may be used as therapeutic targets for the treatment and
CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184 represent the ovarian tumor marker genes of
CC the invention.
XX
SQ Sequence 399 BP; 91 A; 100 C; 111 G; 97 T; 0 other;

Alignment Scores:
Pred. No.: 9.7e-08 Length: 399
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 22 Gaps: 1

09-833799-13B (1-57) x ABA83108 (1-399)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
:::|||||
Db 220 GATCCTGTTGACCCCAACCAAGGAGGAAGCTGGGAAGTCCCGAGTACTTAT 279
:::|||||
Qy 21 IleArgCysAlaMetLeuAspProCysArgCysLeuLysAspThrAspCysProGly 40
:::|||||
Db 280 GGCCAAATGTTTGATCTTAAACCCCAATTTCTGTGATGGTGGCCAGTCAACCGT 339
:::|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
:::|||||

Db 340 GACTTGAAGTGTTCGATGGCATGTGTGGAAATCCGCTTCCCT 387
:::|||||
RESULT 35
AAX16272
ID AAX16272 standard; DNA; 420 BP.
XX
AC AAX16272;
XX
XX 19-APR-1999 (first entry)
XX
XX DNA sequence of ompA-tc-met-SLPI.
XX
XX Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW acute leukemia; ds.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX US5871956-A.
XX
XX 16-FEB-1999.
XX
XX 22-JUL-1994; 94US-0279056.
XX
XX 30-MAR-1987; 87US-0031846.
PR 05-DEC-1984; 84US-0678822.
XX
XX 29-JUL-1986; 86US-0890526.
PR 03-SEP-1986; 86US-0903471.
XX
XX 06-AUG-1990; 90US-0563832.
PR 22-JUL-1994; 94US-0279056.
XX
XX (AMGE-) AMGEN INC.
XX
XX Bandopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
PI WPI; 1999-166640/14.
XX
XX New DNA sequence encoding mammalian serine protease inhibitor -
PT useful for recombinantly producing inhibitors with different
PT specificities and treating diseases such as emphysema, arthritis,
PT muscular dystrophy, and tumour invasion
XX
XX Example 3; Column 37-40; 37pp; English.
XX
XX The present invention describes a DNA sequence (A) encoding an analogue
CC of a mammalian serine protease inhibitor (B). The DNA sequences and
CC recombinant methods allow manufacture of a class of inhibitors of e.g.
CC cathepsin G, elastase, and trypsin, with different specificities. The
CC recombinant serine protease product can be directed to act
CC intracellularly or extracellularly and is useful in treating conditions
CC caused by a disturbance in the native protease/protease inhibitor
CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
CC allows (B) to be recombinantly produced in sufficient quantities and
CC purities so as to provide economical pharmaceutical compositions. (B) is
CC resistant to heat, acid, and a variety of proteolytic enzymes, is
CC thermodynamically stable in extracellular conditions, and exhibits a
CC high degree of self assembly forming an active tertiary structure in the
CC absence of biochemical stimuli. The present sequence represents the DNA
CC sequence of ompA-tc-met-SLPI.
XX
SQ Sequence 420 BP; 123 A; 98 C; 112 G; 87 T; 0 other;

Alignment Scores:
Pred. No.: 1.03e-07 Length: 420
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 20 Gaps: 1

09-833799-13B (1-57) x AAX16272 (1-420)

QY 3 GluProValIysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 245 GATCCGGTTGATACCCGGAACCGGACTCTGCGAAACCGGGTAAATGCCCGTAACCTAT 304
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 305 GGCACGTGCTGATGCTGAACCGCGCAACTTCTGCGAAATGACGGCCAGTGTAAACGA 364
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 365 GATCTGAATGCTGTATGGTATGTCGGCAAACTCTTGTTGTTCCCG 412

RESULT 36

AAC97624
 ID AAC97624 standard; DNA; 436 BP.

XX AC AAC97624;

XX DT 27-FEB-2001 (first entry)

XX DE DNA encoding ompA-tc-met-SLPI.

XX KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;

XX KW protease mediated tissue destruction; emphysema; glomerulonephritis;

XX KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;

XX KW elastase; ds.

XX OS Synthetic.

XX PN US6132990-A.

XX PD 17-OCT-2000.

XX PF 07-JUN-1991; 91US-0712354.

XX PR 03-JAN-1989; 89US-0293042.

XX PR 06-DEC-1984; 84US-0678822.

XX PR 02-DEC-1985; 85US-0803471.

XX PR 29-JUL-1986; 86US-0890526.

XX PR 30-MAR-1987; 87US-0031846.

XX PR 04-AUG-1987; 87US-0082962.

XX PA (AMGE-) AMGEN BOULDER INC.

XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX DR WPI; 2000-678667/66.

XX PT New serine protease inhibitors and DNA sequences for treating a

XX PT protease-mediated condition or tissue destruction e.g. emphysema or

XX PT tumor invasion and for recombinant production of inhibitors

XX PS Example 3; Column 39-40; 47pp; English.

XX CC This invention relates to new purified and isolated mammalian serine

XX CC protease inhibitor proteins which comprise at least 8 cysteine residues

XX CC and no more than 107 amino acids. The protease inhibitors are capable of

XX CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,

XX CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the

XX CC protease inhibitors of the invention, and include oligonucleotide

XX CC sequences used in the isolation and characterisation of the proteins.

XX CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the

XX CC construction of DNA encoding the protease inhibitors. Peptide sequences

XX CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and

XX CC various other peptides used in the isolation of the protease inhibitors.

XX CC The protease inhibitors have cytostatic and anti-inflammatory activity.

XX CC The serine protease inhibitor protein is useful for treating a

XX CC protease-mediated condition, which includes protease mediated tissue

XX CC destruction, e.g. emphysema, arthritis, glomerulonephritis,

XX CC periodontitis, muscular dystrophy or tumour invasion. It is also useful

XX CC for modulating protease activity. The DNAs are useful for producing the

CC proteins, especially by recombinant methods.

XX SQ Sequence 436 BP; 128 A; 102 C; 116 G; 90 T; 0 other;

Alignment Scores:

Pred. No.: 1.08e-07 Length: 436
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97624 (1-436)

QY 3 GluProValIysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 245 GATCCGGTTGATACCCGGAACCGGACTCTGCGAAACCGGGTAAATGCCCGTAACCTAT 304
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 305 GGCACGTGCTGATGCTGAACCGCGCAACTTCTGCGAAATGACGGCCAGTGTAAACGA 364
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 365 GATCTGAATGCTGTATGGTATGTCGGCAAACTCTTGTTGTTCCCG 412

RESULT 37

AAI67583

ID AAI67583 standard; DNA; 436 BP.

XX AC AAI67583;

XX DT 11-FEB-2002 (first entry)

XX DE DNA sequence coding for ompA-tc-met-SLPI.

XX KW Serine protease inhibitor protein; recombinant; leukocyte elastase;

XX KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.

XX OS Homo sapiens.

XX PN US6291662-B1.

XX PD 18-SEP-2001.

XX PF 22-SEP-1998; 98US-0158085.

XX PR 30-MAR-1987; 87US-0031846.

XX PR 06-AUG-1990; 90US-0563832.

XX PR 22-JUL-1994; 94US-0279056.

XX PR 05-DEC-1984; 84US-0678222.

XX PR 02-DEC-1985; 85US-0803471.

XX PR 29-JUL-1986; 86US-0890526.

XX PA (AMGE-) AMGEN INC.

XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX DR WPI; 2001-637974/73.

XX PT New DNA sequences, useful in recombinant DNA techniques for directing

XX PT the production of a serine protease inhibitor protein, e.g. leukocyte

XX PT elastase or trypsin

XX PS Example 3; Column 37-38; 37pp; English.

XX CC The invention relates to a DNA sequence encoding an analog of a mammalian

XX CC serine protease inhibitor protein. The analog comprises at least eight

XX CC cysteine residues and possesses serine protease inhibitor activity, and

XX CC at least one operational element consisting of a promoter, an operator,

XX CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site,

XX CC or a terminator codon. The DNA is useful in recombinant DNA techniques

XX CC for directing the production of a serine protease inhibitor protein,

CC of a mammalian serine protease inhibitor (B). The DNA sequences and

DR WPI; 2000-678667/66.
 XX
 PT New serine protease inhibitors and DNA sequences for treating a
 PT protease-mediated condition or tissue destruction e.g. emphysema or
 PT tumor invasion and for recombinant production of inhibitors
 XX
 XX Example 3; Column 37-38; 47pp; English.
 PS
 CC This invention relates to new purified and isolated mammalian serine
 CC protease inhibitor proteins which comprise at least 8 cysteine residues
 CC and no more than 107 amino acids. The protease inhibitors are capable of
 CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC97582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The protease inhibitors have cytostatic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.
 XX
 SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Alignment Scores:
 Pred. No.: 1.16e-07 Length: 460
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97622 (1-460)

QY 3 GluProValIysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 269 GATCCGGTTGATACCCCGAACCGGACTCTCGAAACCGGGTAATGCCCGTAACCTAT 328
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 329 GGCCAGTGTCTGATGCTGAACCGCGCAACTTCTCGAAATGCGACGCCAGTGTAACGA 388
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 389 GATCTGAAATGCTGTATGGTATGTGCGGCAAAATCTTGTGTTTCCCG 436

RESULT 40
 AA167582
 ID AA167582 standard; DNA; 460 BP.
 XX
 AC AA167582;

DT 11-FEB-2002 (first entry)
 DE DNA sequence coding for ompA SLPI.

KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
 KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.

XX Homo sapiens.
 OS US6291662-B1.
 PN
 XX 18-SEP-2001.
 PD
 XX 22-SEP-1998; 98US-0158085.
 PF
 XX 30-MAR-1987; 87US-0031846.
 PR

PR 06-AUG-1990; 90US-0563832.
 PR 22-JUL-1994; 94US-0279056.
 PR 05-DEC-1984; 84US-0678222.
 PR 02-DEC-1985; 85US-0803471.
 PR 29-JUL-1986; 86US-0890526.
 XX
 XX (AMGE-) AMGEN INC.
 FA
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 PI WPI; 2001-637974/73.
 XX

DR New DNA sequences, useful in recombinant DNA techniques for directing
 PT the production of a serine protease inhibitor protein, e.g. leukocyte
 PT elastase or trypsin
 XX
 PS Example 3; Column 36; 37pp; English.

XX The invention relates to a DNA sequence encoding an analog of a mammalian
 CC serine protease inhibitor protein. The analog comprises at least eight
 CC cysteine residues and possesses serine protease inhibitor activity, and
 CC at least one operational element consisting of a promoter, an operator,
 CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site,
 CC or a terminator codon. The DNA is useful in recombinant DNA techniques
 CC for directing the production of a serine protease inhibitor protein, e.g.
 CC e.g. leukocyte elastase or trypsin. The present sequence represents the
 CC DNA sequence coding for ompA-secretory leukocyte protease inhibitor
 CC (SLPI).

SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Alignment Scores:
 Pred. No.: 1.16e-07 Length: 460
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 22 Gaps: 1

09-833799-13B (1-57) x AA167582 (1-460)

QY 3 GluProValIysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 269 GATCCGGTTGATACCCCGAACCGGACTCTCGAAACCGGGTAATGCCCGTAACCTAT 328
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 329 GGCCAGTGTCTGATGCTGAACCGCGCAACTTCTCGAAATGCGACGCCAGTGTAACGA 388
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 389 GATCTGAAATGCTGTATGGTATGTGCGGCAAAATCTTGTGTTTCCCG 436

RESULT 41
 AA167582
 ID AA167582 standard; cDNA; 498 BP.
 XX
 AC AA167582;

DT 21-FEB-1991 (first entry)

DE Sequence encoding protein with the biological activity of HUSI
 DE (human seminal plasma inhibitor) type I inhibitors encoded on pRH31.

XX Bronchitis therapy; cervix inflammation; hyperfibrinolysis; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
 XX 25..297
 FT CDS /tag= a
 XX

PN DE3600571-A.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

RESULT 46

AAS17499
 ID AAS17499 standard; cDNA; 594 BP.
 AC AAS17499;
 XX
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX
 DE Human cDNA encoding an Antileukoprotease.
 XX
 KW Human; ss; antileukoprotease; antiinflammatory; antiasthmatic;
 KW anti allergic; inflammatory lung disorder; cancer; thyroid tumour;
 KW emphysema; asthma; bronchitis; allergy; ovarian carcinoma;
 KW renal cell carcinoma;
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 19..417
 FT /*tag= a
 FT /product= "Antileukoprotease"
 XX
 FT WO200190421-A2.
 FN 29-NOV-2001.
 PD
 XX
 XX 25-MAY-2001; 2001WO-US17211.
 XX
 PR 25-MAY-2000; 2000US-207104P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Rastelli L, Smithson G;
 PI
 XX
 DR WPI; 2002-106211/14.
 DR P-PSDB; AAU11763.
 XX
 PT Diagnosing inflammatory lung disorders by comparing expression of
 PT nucleic acid encoding antileukoprotease polypeptide in mammal-derived cell
 PT population to nucleic acid encoding the polypeptide in reference
 PT profile -
 XX
 PS Claim 8; Page 19; 36pp; English.
 XX
 CC The invention relates to diagnosing an inflammatory lung disorder in a
 CC mammal, comprising comparing the expression of a nucleic acid encoding
 CC an antileukoprotease polypeptide in mammal-derived cell population,
 CC and expression of a nucleic acid encoding the antileukoprotease in
 CC an inflammation positive or negative reference profile. Also
 CC disclosed are a method for identifying a compound that inhibits lung
 CC inflammation comprising providing a cell expressing antileukoprotease,
 CC contacting the cell with a test compound and measuring the expression of
 CC antileukoprotease in the cell, where a decrease in expression in the
 CC presence of the test compound compared to that in the absence of the test
 CC compound indicates that test compound inhibits lung inflammation.
 CC Also disclosed is assessing prognosis of a mammal with a cancer or
 CC a thyroid tumour, comprising measuring the expression of a nucleic acid
 CC encoding the antileukoprotease in a mammal-derived cell population
 CC and comparing expression of the nucleic acid to the expression of a
 CC nucleic acid encoding the antileukoprotease in a cancer reference
 CC profile, where a substantial similarity between the expression of the
 CC nucleic acid sequence in mammal-derived cell population and the cancer
 CC reference profile indicates an adverse prognosis of the mammal. The
 CC method is used for diagnosing an inflammatory lung disorder such as
 CC emphysema, asthma, bronchitis, or allergy in a mammal. The compound is
 CC useful for treating or preventing an inflammatory disease in a mammal,
 CC preferably human, by inhibiting antileukoprotease. The compound employed
 CC in the method binds to an antileukoprotease nucleic acid, and is a
 CC antileukoprotease antibody, or antileukoprotease antisense nucleic acid.
 CC The methods are useful for assessing the prognosis of a mammal
 CC with thyroid carcinoma, ovarian carcinoma or renal cell carcinoma.
 CC The present sequence is a cDNA encoding a human antileukoprotease.
 XX
 SQ Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 other;

Alignment Scores:
 Pred. No.: 1.58e-07 Length: 594
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 24 Gaps: 1
 09-833799-13B (1-57) x AAS17499 (1-594)
 QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 238 GATCCTGTTGACACCCCAAAACCAAGAGGAGGAGCTGGGAAGTGCCTGACTTAT 297
 QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 298 GGCCAAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGCAGTGCAAGCGT 357
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 358 GACTTGAAGTGTTCATGGCATGTGTGGGAAATCCTGCTTCCCT 405
 RESULT 47
 AAZ77546
 ID AAZ77546 standard; cDNA; 599 BP.
 XX
 AC AAZ77546;
 XX
 DT 10-APR-2000 (first entry)
 XX
 DE Human ovarian tumor cDNA library derived EST fragment 97.
 KW
 KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;
 KW gene therapy; treatment; ss.
 XX
 OS Homo sapiens.
 XX
 FN DE19817557-A1.
 XX
 PD 21-OCT-1999.
 XX
 PF 09-APR-1998; 98DE-1017557.
 XX
 PR 09-APR-1998; 98DE-1017557.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX WPI; 1999-591920/51.
 DR
 XX New nucleic acid sequences expressed in ovarian, and some other, cancer
 XX tissues, and derived polypeptides, for treatment of ovarian cancer and
 XX identification of therapeutic agents -
 PS Claim 3; Page 221; 310pp; German.
 XX
 CC This invention describes novel nucleic acid (cDNA) sequences (A) which
 CC have anticancer activity and are highly expressed in ovarian tumor
 CC tissue (and some also in testis and breast cancer tissue). The products
 CC of the invention can be used for gene therapy. (A) are used (i) for
 CC recombinant expression of polypeptides (B) and (ii) to isolate complete
 CC genes. (B) are used (i) to identify agents suitable for treatment of
 CC ovarian cancer; (ii) directly for treating this form of cancer
 CC (including expression from gene therapy vectors) and (iii) for generation
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent different
 CC parts of the same unknown gene, distorting the estimated frequency of
 CC occurrence in a particular tissue. AAZ77450-277572 represent the human


```

PD XX 27-JUN-2002.
PF XX
PR XX 18-DEC-2001; 2001WO-US49256.
PR XX
PR XX 18-DEC-2000; 2000US-256699P.
PR XX
PR XX 20-NOV-2001; 2001US-331966P.
PR XX
PA (ARRI-) ARRIVA PHARM INC.
XX XX
PI Barr PJ, Gibson HL, Pemberton P;
XX XX
DR WPI; 2002-500631/53.
DR P-PSDB; AAU99881.
XX XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX XX
PS Example 1; Page 73-73; 134pp; English.
XX XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC SLAP1 fusion protein of the invention.
XX XX
SQ Sequence 1525 BP; 467 A; 286 C; 314 G; 458 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 5, 02e-07 Length: 1525
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 24 Gaps: 1

```

```

09-833799-13B (1-57) x ABK88022 (1-1525)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysPProIleLeu 20
Db 156 GACCAGTTGACACCCCAACCACTAGAGAAAGCCAGGTAGTGCAGTACTTAC 215
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 216 GGTCAATGTTTGTATGTTGAACCCACCAACCTCTGTGAATGGCGGTCAATGAAGA 275
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 276 GACTTGAAGTGTGTATGGTATGTGTGTGAAGTCTGTGTGTTTCCCA 323

```

```

RESULT 50
ABK88025
ID ABK88025 standard; DNA; 1525 BP.
XX
AC ABK88025;
XX
DT 07-OCT-2002 (first entry)
XX

```

```

DE DNA sequence encoding rSLAP1 fusion protein.
XX
XX rSLAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis externa; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW glomerulonephritis; hypertension.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
RBS 6..8
FT /tag= a
FT /standard_name= "Ribosome binding site"
FT 9..1520
FT /tag= b
FT /product= "rSLAP1 fusion protein"
FT 12..1193
FT /tag= c
FT /note= "AAT coding region"
FT 1194..1196
FT /tag= d
FT /note= "linking codon"
FT 1197..1517
FT /tag= e
FT /note= "SLPI coding region"
XX
XX WO200250287-A2.
XX
XX 27-JUN-2002.
XX
XX 18-DEC-2001; 2001WO-US49256.
XX
XX 18-DEC-2000; 2000US-256699P.
XX
XX 20-NOV-2001; 2001US-331966P.
XX
XX (ARRI-) ARRIVA PHARM INC.
XX
XX Barr PJ, Gibson HL, Pemberton P;
XX
XX WPI; 2002-500631/53.
XX
XX P-PSDB; AAU99884.
XX
XX Novel fusion protein useful for inhibiting protease activity associated
XX with a disorder such as emphysema, asthma, comprises a first protease
XX inhibitor comprising alpha 1-antitrypsin and a second protease
XX inhibitor -
XX
XX Example 3; Page 89-90; 134pp; English.
XX
XX This invention relates to a novel fusion protein comprising a first
XX protease inhibitor comprising an alpha1-antitrypsin or its functionally
XX active portion and a second protease inhibitor or its functionally
XX active protein. The fusion proteins of the invention may act as an
XX inhibitor of protease activity. The fusion protein of the invention
XX is useful for inhibiting protease activity associated with a disorder
XX such as emphysema, asthma, chronic obstructive pulmonary disease,
XX cystic fibrosis, otitis media, otitis externa or HIV infection, or
XX for treating an individual suffering from or at risk for a disease or
XX disorder involving unwanted protease activity. The proteins are useful
XX for treating dermatological diseases such as atopic dermatitis, eczema
XX and psoriasis, in inflammatory responses to viral infection, and for
XX treating herpes infection, corneal or epidermal ulceration, chronic
XX non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
XX tumour metastasis and tumour angiogenesis, gastric ulceration,
XX osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
XX bacterial infection, Alzheimer's disease, hypertension and muscular
XX dystrophy. The present sequence represents the DNA encoding the
XX rSLAP1 fusion protein of the invention.
XX

```

SQ Sequence 1525 BP; 467 A; 287 C; 314 G; 457 T; 0 other;

Alignment Scores:
 Pred. No.: 5.02e-07 Length: 1525
 Score: 143.00 Matches: 27
 Percent Similarity: 55.36% Conservative: 4
 Best Local Similarity: 48.21% Mismatches: 23
 Query Match: 43.73% Indels: 2
 DB: 24 Gaps: 1

09-833799-13B (1-57) x ABK88025 (1-1525)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 1341 GACCCAGTTGACACCCCAACCACTAGAGAAAGCCAGTAAAGTGTCCAGTTACTTAC 1400
 Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 1401 GGTCAATGTTTGTGTTGAACCCCAAACTTCTGTGAAATGGACGGTCAATGTAAGAGA 1460
 Qy 41 IleLysLysCysCysGlyGlySerCysGlyMetAlaCysPheValPro 56
 Db 1461 GACTTGAAGTGTGTATGGGTATGTGTGTAAGTCCGTGTTTCCCA 1508

RESULT 51

AAN60465
 ID AAN60465 standard; DNA; 2274 BP.

XX AC AAN60465;

XX DT 01-JAN-1980 (first entry)

XX DE Synthetic sequence capable of directing microbial synthesis of a
 XX secretory leukocyte protease inhibitor.

XX KW Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
 XX trypsin; inhibitor; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT exon 882..1042

XX FT exon /*tag= a

XX FT exon 1458..1607

XX FT exon /*tag= b

XX FN W08603519-A.

XX PD 19-JUN-1986.

XX PF 04-DEC-1985; 85WO-US02385.

XX PR 04-DEC-1985; 85WO-US02385.

XX PR 06-DEC-1984; 84US-0678822.

XX PR 02-DEC-1985; 85US-0803471.

XX PA (SYNE-) SYNERGEN BIOLOG INC.

XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX DR WPI; 1986-169458/26.

XX DR P-PSDB; AAP60565, AAP60566.

XX PT New synthetic DNA sequences for directing microbial synthesis -
 XX for prodn. of single polypeptide chain serine protease inhibitor
 XX having leukocyte elastase and trypsin inhibitory sites

XX PS Disclosure; Page 19; 59pp; English.

XX CC The sequence directs synthesis of a preferred secretory leukocyte
 XX protease-inhibitor. See also AAN60464, AAN60466-69 and AAP60562-
 XX 66.

SQ Sequence 2274 BP; 459 A; 567 C; 602 G; 646 T; 0 other;

Alignment Scores:
 Pred. No.: 1.39e-06 Length: 2274
 Score: 141.00 Matches: 24
 Percent Similarity: 55.10% Conservative: 3
 Best Local Similarity: 48.98% Mismatches: 22
 Query Match: 43.12% Indels: 0
 DB: 7 Gaps: 0

09-833799-13B (1-57) x AAN60465 (1-2274)

Qy 8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn 27
 Db 1454 CCAGCAAGGAGGAGAGCTTGGGAAGTGCAGTATATGGCAATGTTGATGCTTAAC 1513
 Qy 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysGlyGly 47
 Db 1514 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAGTCAAGCGTGACTTGAAGTGTTCATGGGC 1573
 Qy 48 SerCysGlyMetAlaCysPheValPro 56
 Db 1574 ATGTGTGGGAATCTGTGTTCCCT 1600

RESULT 52

AAC97530

ID AAC97530 standard; DNA; 2274 BP.

XX AC AAC97530;

XX DT 27-FEB-2001 (first entry)

XX DE Human genomic DNA encoding SLPI.

XX KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 XX protease mediated tissue destruction; emphysema; glomerulonephritis;
 XX periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 XX elastase; ds.

XX OS Homo sapiens.

XX PN US6132990-A.

XX PD 17-OCT-2000.

XX PF 07-JUN-1991; 91US-0712354.

XX PR 03-JAN-1989; 89US-0293042.

XX PR 06-DEC-1984; 84US-0678822.

XX PR 02-DEC-1985; 85US-0803471.

XX PR 29-JUL-1986; 86US-0890526.

XX PR 30-MAR-1987; 87US-0031846.

XX PR 04-AUG-1987; 87US-0082962.

XX PA (AMGE-) AMGEN BOULDER INC.

XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX DR WPI; 2000-678667/66.

XX DR P-PSDB; AAB53106.

XX PT New serine protease inhibitors and DNA sequences for treating a
 XX protease-mediated condition or tissue destruction e.g. emphysema or
 XX tumor invasion and for recombinant production of inhibitors -

XX PS Disclosure; Column 13-18; 47pp; English.

XX CC This invention relates to new purified and isolated mammalian serine
 XX protease inhibitor proteins which comprise at least 8 cysteine residues
 XX and no more than 107 amino acids. The protease inhibitors are capable of
 XX inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 XX AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 XX protease inhibitors of the invention, and include oligonucleotide

CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The protease inhibitors have cytosolic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX Sequence 2274 BP; 459 A; 567 C; 602 G; 646 T; 0 other;

Alignment Scores:

Pred. No.: 1.39e-06 Length: 2274
 Score: 141.00 Matches: 24
 Percent Similarity: 55.10% Conservative: 3
 Best Local Similarity: 48.98% Mismatches: 22
 Query Match: 43.12% Indels: 0
 DB: 21 Gaps: 0

09-833799-13B (1-57) x AAC97530 (1-2274)

QY 8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn 27
 DB 1454 CCAGCAAGGAGGAGCTGGGAGTGGCCAGTACTTATGGCCAAATGTTGATGCTTAAC 1513
 QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
 DB 1514 CCCCCCAATTTCTGTGAGATGGATGCCAGTGCACAGCGTGACTTGAAGTGTTCATGGGC 1573
 QY 48 SerCysGlyMetAlaCysPheValPro 56
 DB 1574 ATGTGTGGGAAATCCTGCTTCCCT 1600

RESULT 53

AAI67491
 ID AAI67491 standard; DNA; 2274 BP.

XX AAI67491;

XX 11-FEB-2002 (first entry)

XX Recombinant secretory leukocyte protease inhibitor (SLPI) DNA.

XX Serine protease inhibitor protein; recombinant; leukocyte elastase;
 KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 884..1607
 FT /*tag= a
 FT /product= "SLPI"
 FT /note= "contains introns"
 FT exon 884..1039
 FT /*tag= b
 FT intron 1040..1459
 FT /*tag= c
 FT exon 1460..1607
 FT /*tag= d

XX US6291662-B1.

XX 18-SEP-2001.

XX 22-SEP-1998; 98US-0158085.

XX 30-MAR-1987; 87US-0031846.

XX 06-AUG-1990; 90US-0563832.

XX 22-JUL-1994; 94US-0279056.

PR 05-DEC-1984; 84US-0678222.
 PR 02-DEC-1985; 85US-0803471.
 PR 29-JUL-1986; 86US-0890526.
 XX (AMGE-) AMGEN INC.
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 PI WPI; 2001-637974/73.
 DR P-PSDB; AAG65996.
 XX

XX New DNA sequences, useful in recombinant DNA techniques for directing
 PT the production of a serine protease inhibitor protein, e.g. leukocyte
 PT elastase or trypsin

XX Disclosure; Columns 13-18; 37pp; English.

XX The invention relates to a DNA sequence encoding an analog of a mammalian
 CC serine protease inhibitor protein. The analog comprises at least eight
 CC cysteine residues and possesses serine protease inhibitor activity, and
 CC at least one operational element consisting of a promoter, an operator,
 CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
 CC or a terminator codon. The DNA is useful in recombinant DNA techniques
 CC for directing the production of a serine protease inhibitor protein,
 CC e.g. leukocyte elastase or trypsin. The present sequence represents the
 CC DNA that encodes a secretory leukocyte protease inhibitor (SLPI).

XX Sequence 2274 BP; 459 A; 567 C; 602 G; 646 T; 0 other;

Alignment Scores:

Pred. No.: 1.39e-06 Length: 2274
 Score: 141.00 Matches: 24
 Percent Similarity: 55.10% Conservative: 3
 Best Local Similarity: 48.98% Mismatches: 22
 Query Match: 43.12% Indels: 0
 DB: 22 Gaps: 0

09-833799-13B (1-57) x AAI67491 (1-2274)

QY 8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn 27
 DB 1454 CCAGCAAGGAGGAGCTGGGAGTGGCCAGTACTTATGGCCAAATGTTGATGCTTAAC 1513

QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
 DB 1514 CCCCCCAATTTCTGTGAGATGGATGCCAGTGCACAGCGTGACTTGAAGTGTTCATGGGC 1573

QY 48 SerCysGlyMetAlaCysPheValPro 56

DB 1574 ATGTGTGGGAAATCCTGCTTCCCT 1600

RESULT 54

AAQ45442
 ID AAQ45442 standard; DNA; 180 BP.

XX AAQ45442;

XX 26-OCT-1994 (first entry)

XX CLPI.

XX Serine leukocyte protease inhibitor; SLPI; truncated SLPI; CLPI;
 KW retrovirus; retroviral infection; HIV; parotid inhibitor; enzyme;
 KW elastase; cathepsin G; trypsin; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT misc_feature 79..81

FT /*tag= a

XX /transl_except= pos:79..81, aa:Leu

XX W09406454-A.


```

XX PD 31-MAR-1994.
XX PR
XX PF 09-SEP-1993; 93WO-US08486.
XX PP
XX PR 09-SEP-1992; 92US-0943369.
XX PP
XX PA (SYND ) SYNERGEN INC.
XX PS (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Eisenberg S, Thompson RC, Wahl SM;
XX DR WPI; 1994-118153/14.
XX PP P-PSDB; AAR511172.
XX PR Inhibiting retroviral infection with serine leucocyte protease
XX PT inhibitor - esp. for treating or preventing HIV infection, also
XX PT new protease inhibitors and nucleic acids encoding them
XX PS
XX PS Claim 29; Page 17; 21pp; English.
XX CC Retroviral infection is inhibited by admin. of an SLPI, or its
XX CC analogue or deriv. A useful SLPI deriv. is CLPI, a truncated SLPI
XX CC mol. having only the last 60 amino acids of the native parotid
XX CC inhibitor. SLPI acts by blocking a host cell enzyme essential for
XX CC retroviral infection; it is a powerful inhibitor of human leucocyte
XX CC elastase, cathepsin G and human trypsin. The inhibitor provides an
XX CC effective and relatively safe method for treating HIV infections.
XX SQ Sequence 180 BP; 42 A; 45 C; 50 G; 43 T; 0 other;

Alignment Scores:
Pred. No.: 1.06e-07 Length: 180
Score: 139.00 Matches: 26
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 46.43% Mismatches: 24
Query Match: 42.51% Indels: 2
DB: 15 Gaps: 1

09-833799-13B (1-57) x AAQ45442 (1-180)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 4 GATCTGTGTGACACCCCAACCAACAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 63
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 64 GGCCAATGTTTGATGCCCTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 123
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 124 GACTTGAAGTGTGATGGCATGTGTGGGAAATCCTCGCTTCCCT 171

RESULT 55
AAZ88483
ID AAZ88483 standard; DNA; 180 BP.
XX AC
XX AC AAZ88483;
XX DT
XX DT 12-MAY-2000 (first entry)
XX DE
XX DE Truncated SLPI encoding nucleotide sequence CLPI SEQ ID NO:2.
XX KW Secretory leukocyte protease inhibitor; retroviral; infection; HIV;
XX KW AIDS; SLPI; CLPI; parotid inhibitor; antiviral; anti-HIV; ss.
XX OS Homo sapiens.
XX XX
XX PN US6017880-A.
XX PD
XX PD 25-JAN-2000.
XX PF
XX PF 07-JUN-1995; 95US-0483503.

```

```

XX PR 09-SEP-1992; 92US-0943369.
XX PR 09-SEP-1993; 93WO-US08486.
XX PR 09-MAR-1994; 94US-0209040.
XX XX
XX PA (AMGE-) AMGEN INC.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Dripps DJ, Thompson RC, Wahl SM, Eisenberg S;
XX DR WPI; 2000-170378/15.
XX PP P-PSDB; AAY79951.
XX PR Inhibiting retrovirus infection using secretory leukocyte protease
XX PT inhibitors for treating diseases such as acquired immunodeficiency
XX PT syndrome -
XX PS
XX PS Disclosure; Column 3-5; 18pp; English.
XX CC The present invention describes a method for inhibiting, in vitro,
XX CC retrovirus infection of CD4+ cells comprising contacting the cells with
XX CC a secretory leukocyte protease inhibitor (SLPI) having the amino acid
XX CC sequence of a naturally occurring SLPI or a substitution analogue having
XX CC 107 amino acid residues. SLPI inhibits enzymes whose function are
XX CC necessary for retroviral infection of the cell. SLPI is a potent
XX CC inhibitor of elastase, trypsin, cathepsin G and chymotrypsin. The method
XX CC is useful for treating retrovirus infection in a patient and associated
XX CC diseases such as acquired immunodeficiency syndrome (AIDS). The present
XX CC sequence encodes a truncated version of SLPI designated CLPI.
XX SQ Sequence 180 BP; 42 A; 45 C; 50 G; 43 T; 0 other;

Alignment Scores:
Pred. No.: 1.06e-07 Length: 180
Score: 139.00 Matches: 26
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 46.43% Mismatches: 24
Query Match: 42.51% Indels: 2
DB: 21 Gaps: 1

09-833799-13B (1-57) x AAZ88483 (1-180)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 4 GATCTGTGTGACACCCCAACCAACAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 63
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 64 GGCCAATGTTTGATGCCCTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 123
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 124 GACTTGAAGTGTGATGGCATGTGTGGGAAATCCTCGCTTCCCT 171

RESULT 56
AAZ37665
ID AAA37665 standard; DNA; 1010 BP.
XX AC
XX AC AAA37665;
XX DT
XX DT 24-OCT-2000 (first entry)
XX DE
XX DE Human peptidase, HPEP-9 coding sequence.
XX KW Human; peptidase; cell proliferative disorder; arteriosclerosis;
XX KW psoriasis; myelofibrosis; cancer; autoimmune disorder; Crohn's disease;
XX KW inflammatory disorder; AIDS; anaemia; allergy; asthma; atherosclerosis;
XX KW Grave's disease; multiple sclerosis; scleroderma; infection; diabetes;
XX KW metabolic disorder; Addison's disease; cystic fibrosis; diagnosis;
XX KW glycogen storage disease; obesity; therapy; HPEP-9; ds.
XX OS
XX OS Homo sapiens.

```

```

FH Key      Location/Qualifiers
FT CDS      90..461
FT          /*tag= a
FT          /product= HPEP-9
XX
PN WO200042201-A2.
XX
XX 20-JUL-2000.
XX
XX 11-JAN-2000; 2000MO-US00641.
XX
XX 11-JAN-1999; 99US-0172247.
XX
XX 03-MAY-1999; 99US-0132253.
XX
XX 27-MAY-1999; 99US-0136653.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Hillman JL, Tang YT, Azimzai Y, Baughn MR, Lal P;
XX Yue H, Lu DAM;
XX
XX WPI: 2000-482832/42.
XX P-PSDB; AAY90292.
XX
XX An isolated polypeptide for diagnosis, prevention and treatment of
XX cell proliferative, autoimmune/ inflammatory and metabolic disorders
XX comprises a sequence encoding a human peptidase .
XX
XX Claim 4; Page 118; 131pp; English.
XX
XX This sequence encodes the human peptidase, designated HPEP-9. The
XX invention relates to 18 human peptidases designated HPEP-1 to HPEP-18,
XX respectively. The peptidases can be used for treating a disease or
XX condition associated with decreased expression or over expression of
XX functional human peptidases. The diseases that can be diagnosed or
XX prevented and treated include cell proliferative disorders (such as
XX arteriosclerosis, psoriasis, myelofibrosis, and cancers),
XX autoimmune/inflammatory disorders (such as AIDS, anaemia, allergies,
XX Crohn's disease, asthma, atherosclerosis, Grave's disease, multiple
XX sclerosis, and scleroderma), infections, and metabolic disorders (such as
XX Addison's disease, diabetes, cystic fibrosis, glycogen storage diseases
XX and obesity).
XX
XX Sequence 1010 BP; 223 A; 308 C; 243 G; 236 T; 0 other;

Alignment Scores:
Pred. No.:      5.4e-05      Length:      1010
Score:          123.50      Matches:      25
Percent Similarity: 61.22%      Conservative: 5
Best Local Similarity: 51.02%      Mismatches: 16
Query Match:      37.77%      Indels:      3
DB:               21         Gaps:        2

09-833799-13B (1-57) x AAA37665 (1-1010)

QY 9 ValSerThrLysProGlySerCysProIleLeuIleLeuIleArgCysAlaMetLeuAsnPro 28
Db 309 GTCTCTGTGAAGCTGGGAGCTGCCAGAGGACCAACTGCGCTGC-----CTCAGCCCC 362
QY 29 ProAsnArg---CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
Db 363 ATGAACACCACTGTGTGTACAAAGGACTCAGACTGCTCGGCAAAAGCGATGCTGCCACAGC 422
QY 48 SerCysGlyMetAlaCysPheValPro 56
Db 423 GCCTGCGGCGGGATTGCCGGGATCCT 449

RESULT 57
AAD38698
ID AAD38698 standard; cDNA; 1013 BP.
XX
XX AAD38698;
XX
XX 23-SEP-2002 (first entry)

```

```

XX
XX Human LP229 secreted protein encoding cDNA.
XX
XX Human; secreted protein; atherosclerosis; Alzheimer's disease; LP229;
XX diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
XX rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
XX reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
XX gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
XX chromosome 20q12-13.2; gene; ss.
XX
XX Homo sapiens.
XX
XX Key      Location/Qualifiers
XX CDS      90..461
XX          /*tag= a
XX          /product= "Human LP229 secreted protein"
XX
XX sig_peptide 90..143
XX          /*tag= b
XX
XX mat_peptide 144..458
XX          /*tag= c
XX          /product= "Mature human LP229 secreted protein"
XX
XX WO200226801-A2.
XX
XX 04-APR-2002.
XX
XX 14-SEP-2001; 2001WO-US26026.
XX
XX 28-SEP-2000; 2000US-236088P.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Su EW, Wang H;
XX
XX WPI: 2002-471259/50.
XX P-PSDB; AAE23983.
XX
XX Novel proteins and polynucleotides of secreted proteins useful for
XX treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
XX diabetic retinopathy, arteriosclerosis, ischaemia or reperfusion injury
XX
XX Claim 1; Page 138-139; 145pp; English.
XX
XX The invention relates to human secreted polypeptides designated LP095,
XX LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic
XX acid molecules encoding such polypeptides. Novel secreted proteins of
XX the invention are used for treating diseases such as atherosclerosis,
XX Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
XX arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
XX combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
XX reperfusion injury, neoplasms and cancer especially liver cancer. They
XX are also used for wound healing. Polynucleotides of the invention can
XX be used to generate transgenic animals or knock out animals, which in
XX turn, are useful in the development and screening of therapeutically
XX useful reagents for use in the treatment of diseases associated with
XX LP polypeptide associated activity. They are also used in gene therapy.
XX The present sequence is human LP229 secreted protein encoding cDNA.
XX LP229 gene is located on chromosome 20q12-13.2.
XX
XX Sequence 1013 BP; 226 A; 308 C; 243 G; 236 T; 0 other;

Alignment Scores:
Pred. No.:      5.42e-05      Length:      1013
Score:          123.50      Matches:      25
Percent Similarity: 61.22%      Conservative: 5
Best Local Similarity: 51.02%      Mismatches: 16
Query Match:      37.77%      Indels:      3
DB:               24         Gaps:        2

09-833799-13B (1-57) x AAD38698 (1-1013)

QY 9 ValSerThrLysProGlySerCysProIleLeuIleLeuIleArgCysAlaMetLeuAsnPro 28

```


XX DT 26-NOV-2001 (first entry)
 XX DE Nucleotide sequence of a human secreted polypeptide.
 XX Human; secreted polypeptide; nervous disease; muscular disease; tumour;
 KW gastrointestinal ulceration; spinal cord disease; trachea disease;
 KW thyroid gland disease; ovary disease; prostate disease; thymus disease;
 KW renal gland disease; small intestine disease; colon disease;
 KW lymph node disease; muscular system disease; thymus disease;
 KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
 KW myocardial infarction; angiotensin; liver disease; coagulation disorder;
 KW microbial infection; immune disorder; inflammation; transplant rejection;
 KW bone thickness; bone density; ferroxidase loss; apoptosis;
 KW vascular smooth cell proliferation; vaccine; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 1..732
 FT /tag= a
 FT /product= "secreted polypeptide"
 XX W0200166690-A2.
 XX PD 13-SEP-2001.
 XX PF 05-MAR-2001; 2001WO-US07143.
 XX PR 06-MAR-2000; 2000US-0187107.
 XX PR 13-MAR-2000; 2000US-0189916.
 XX PR 03-OCT-2000; 2000US-0236874.
 XX PR 03-OCT-2000; 2000US-0237846.
 XX (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 XX PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
 XX WPI; 2001-570768/64.
 DR P-PSDB; AAG67518.
 XX Novel isolated secreted polypeptide useful for treating nervous and
 PT muscular diseases, gastrointestinal ulceration, coagulation and immune
 PT disorders, microbial diseases, inflammation and transplant rejection -
 XX Claim 2; Page 47; 102pp; English.
 CC The present sequence encodes a human secreted polypeptide. The
 CC secreted polypeptides and polynucleotides are useful for treating
 CC nervous and muscular diseases, for inhibiting tumour formation and
 CC metastasis, for treating gastrointestinal ulceration, for preventing
 CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,
 CC renal gland, small intestine, heart, trachea, thymus, lymph node,
 CC muscular system and colon, for treating lipase deficiency in cystic
 CC fibrosis and pancreatitis, for treating undesirable clot formation
 CC such as myocardial infarction, during angioplasty and all surgical
 CC procedures that require decreased blood clot formation, for treating
 CC liver diseases, coagulation disorders and microbial diseases, for
 CC treating immune disorders, for treating inflammation and transplant
 CC rejection, for enhancing bone thickness and increasing bone density,
 CC for reducing the loss of essential ferroxidases, for suppressing
 CC apoptosis, and for regulating vascular smooth cell proliferation. They
 CC may also be used as vaccines.
 XX SQ Sequence 732 BP; 149 A; 212 C; 215 G; 156 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.000138 Length: 732
 Score: 118.50 Matches: 24
 Percent Similarity: 61.22% Conservative: 6
 Best Local Similarity: 48.98% Mismatches: 16
 Query Match: 36.24% Indels: 3

DB: 22 Gaps: 2
 09-833799-13B (1-57) x AAH78210 (1-732)
 Qy 9 ValSerThrIysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnPro 28
 Db 530 GTTCAGTGAAGTGGCAGCTGCCAGAGGACCACTGGCTGC-----CTCAGCCCC 583
 Qy 29 ProAsnArg--CysLeuLysAspThrAspCysProGlyIleLysCysCysGluGly 47
 Db 584 ATGAACACCTGTGTACACAGGACTCAGACTGCTCGGGCAAAAGCGATGCTGCCACAGC 643
 Qy 48 SerCysGlyMetAlaCysPheValPro 56
 Db 644 GCCTGCGGGGGATTGCGCGGATCCT 670
 RESULT 62
 AAL07251
 ID AAL07251 standard; DNA; 5690 BP.
 XX AC AAL07251;
 XX DT 21-NOV-2001 (first entry)
 XX DE Human reproductive system related antigen DNA SEQ ID NO: 9939.
 KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ds.
 XX OS Homo sapiens.
 XX WO200155320-A2.
 XX PD 02-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01339.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 28-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0231242.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
DR
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Disclosure; SEQ ID NO 9939; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
SQ Sequence 5690 BP; 1336 A; 1453 C; 1505 G; 1396 T; 0 other;

Alignment Scores:
Pred. No.: 0.0017 Length: 5690
Score: 118.50 Matches: 24
Percent Similarity: 61.22% Conservative: 6
Best Local Similarity: 48.98% Mismatches: 16
Query Match: 36.24% Indels: 3
DB: 22 Gaps: 2

09-833799-13B (1-57) x AAL07251 (1-5690)
QY 9 ValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnPro 28
Db 4606 GTTCAGTGAAGCTGGGAGCTGCCAGAGACCACTGGCGTGC-----CTCAGCCCC 4659
QY 29 ProAsnArg---CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
Db 4660 ATGAACCACTGTGTCAAGAGACTCAGACTGCTCGGCAAAAGCGATGTCACAGC 4719
QY 48 SerCysGlyMetAlaCysPheValPro 56
Db 4720 GCCTGGGGCGGGATTGCGGGATCCT 4746

RESULT 63
ABL98799
ID ABL98799 standard; DNA; 5690 BP.
XX

AC ABL98799;
XX DT 21-JUN-2002 (first entry)
XX DE Human testicular antigen encoding DNA fragment SEQ ID NO: 3451.
XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX OS Homo sapiens.
XX PN WO20015317-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01329.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0223267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251899.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2000US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483232/52.
 XX
 PT Nucleic acids encoding 973 human testicular antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating testicular cancer -
 PS Disclosure; SEQ ID NO 3451; 766pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 973
 CC human testicular antigens, and fragments of their genomic sequences. The
 CC sequences can be used in the treatment of cardiovascular, urinary system,
 CC reproductive system, immune, respiratory, neurological and
 CC gastrointestinal disorders, infections, and particularly cancer,
 CC especially testicular cancers. The present sequence is a DNA encoding a
 CC protein fragment of the invention.
 XX
 SQ Sequence 5690 BP; 1336 A; 1453 C; 1505 G; 1396 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0017 Length: 5690
 Score: 118.50 Matches: 24
 Percent Similarity: 61.22% Conservative: 6
 Best Local Similarity: 48.98% Mismatches: 16
 Query Match: 36.24% Indels: 3
 DB: 23 Gaps: 2

09-833799-13B (1-57) x ABL98799 (1-5690)

Qy 9 ValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnPro 28
 Db 4606 GTTCAGTGAAGCTGGGAGCTGCCAGAGGACCAACTGCGCTGC-----CTCAGCCCC 4659
 Qy 29 ProAsnArg---CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
 Db 4660 ATGAACCACTGTGTCACAGGACTCAGACTGCTCGGGCAAAAGCGATCTGCCACAGC 4719
 Qy 48 SerCysGlyMetAlaCysPheValPro 56
 Db 4720 GCCTCGCGGGCGGATTGCGGGATCCT 4746

RESULT 64
 AAN60469
 ID AAN60469 standard; DNA; 872 BP.
 XX
 AC AAN60469;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE ompA-tc-met-secretory leukocyte protease-inhibitor from a phage M13mpl8
 DE construct.
 XX
 KW Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
 KW trypsin-inhibitor; ompA-tc-met-SLPI; ds.
 XX
 OS Synthetic.
 XX
 PN WO8603519-A.
 XX
 PD 19-JUN-1986.
 XX

PF 04-DEC-1985; 85WO-US02385.
 XX
 PR 04-DEC-1985; 85WO-US02385.
 PR 06-DEC-1984; 84US-0678922.
 PR 02-DEC-1985; 85US-0803471.
 XX
 PA (SYNE-) SYNERGEN BIOLOG INC.
 XX
 PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX
 DR WPI; 1986-169458/26.
 XX
 PT New synthetic DNA sequences for directing microbial synthesis -
 PT for prodn. of single poly:peptide chain serine protease inhibitor
 PT having leukocyte elastase and trypsin inhibitory sites
 XX
 PS Disclosure; Page 46; 59pp; English.
 XX
 CC The sequence directs synthesis of a secretory leukocyte protease-
 CC inhibitor. See also AAN60463-68, and AAP60562-66.
 XX
 SQ Sequence 872 BP; 218 A; 218 C; 218 G; 218 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000331 Length: 872
 Score: 116.00 Matches: 29
 Percent Similarity: 36.46% Conservative: 6
 Best Local Similarity: 30.21% Mismatches: 19
 Query Match: 35.47% Indels: 42
 DB: 7 Gaps: 4

09-833799-13B (1-57) x AAN60469 (1-872)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
 Db 485 GATCGGTTGATACCCCGAACCCGACTCGTGGAAACCGGGTAAATGCCGGTAAACGGAC 544
 Qy 21 Ile-----GlyIleLys----- 21
 Db 545 CTAGGCCAACCTATGGGCTTGGGCTGAGCAGCTTTTGGCCCATTTACGGGCCATTGCTAT 604
 Qy 22 ---ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysPro--- 39
 Db 605 GCCCAGTGTCTGATGCTGAACCCCGCGAACTTTCGGAATGACGCCGACGCTGAAGATA 664
 Qy 40 -----GlyIleLys----- 42
 Db 665 CCGGTCACAGACTACGACTTGGGCGGCTTGAAGACGCTTTACCTGCGGTCACATTACGA 724
 Qy 43 -----LysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 725 GATCTGAAATGCTGATGGGTATGTGGCGCAATCTTGTGTTTCCCG 772

RESULT 65
 AAN60468
 ID AAN60468 standard; DNA; 920 BP.
 XX
 AC AAN60468;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE ompA-secretory leukocyte protease-inhibitor from a phage M13mpl8
 DE construct.
 XX
 KW Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
 KW trypsin-inhibitor; ompA-SLPI; ds.
 XX
 OS Synthetic.
 XX
 PN WO8603519-A.
 XX
 PD 19-JUN-1986.
 XX


```

PF 04-DEC-1985; 85WO-US02385.
XX
PR 04-DEC-1985; 85WO-US02385.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
XX
XX (SYNE-) SYNERGEN BIOLOG INC.
XX
XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
PI WPI; 1986-169458/26.
XX
XX New synthetic DNA sequences for directing microbial synthesis -
PT for prodn. of single poly-peptide chain serine protease inhibitor
PT having leukocyte elastase and trypsin inhibitory sites
XX
XX Disclosure; Page 46; 59pp; English.
XX
XX The sequence directs synthesis of a secretory leukocyte protease-
CC inhibitor. See also AAN60463-67, AAN60469, and AAP60562-66.
CC
XX Sequence 920 BP; 221 A; 239 C; 238 G; 221 T; 1 other;
SQ
Alignment Scores:
Pred. No.: 0.000354 Length: 920
Score: 116.00 Matches: 22
Percent Similarity: 40.30% Conservative: 5
Best Local Similarity: 32.84% Mismatches: 20
Query Match: 35.47% Indels: 20
DB: 7 Gaps: 1

09-833799-13B (1-57) x AAN60468 (1-920)
QY 10 SerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsnProPro 29
Db 596 GCAGCAAAACCGGGTAAATGCCCGGTAACTATGGCCAGTGTCTGATGCTGAACCGCGG 655
QY 30 Asn----- 30
Db 656 AACTTTTGGCCATTACGGGCCATTGGATACCGGTACAGACTACGACTTGGCGGC 715
QY 31 ---ArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluCysCys 49
Db 716 TTGAACCTGCAATGACCGCGGAGTAAACGAGATCTTGAAATGCTGTGGGTATGTC 775
QY 50 GlyMetAlaCysPheValPro 56
Db 776 GGCAAGACGCTTTACCTGCGG 796

RESULT 66
AAx80909
ID AAx80909 standard; cdna; 478 BP.
XX
XX AAx80909;
XX
DT 03-NOV-1999 (first entry)
XX
XX Human cdna clone HKABR62.
XX
XX Human cdna clone HKABR62; secreted protein; serine protease inhibitor;
XX serine protease; Mus musculus secretory leukocyte protease inhibitor;
XX extracellular matrix degradation; multiple sclerosis; cancer; arthritis;
XX inflammation; immune system disorder; neurodegenerative disorder; serpin;
XX Kallmann's syndrome; Down's syndrome; Alzheimer's; fusion protein;
XX galactorrhea; hypogonadism; somatostatin; protein purification; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 19..249
FT CDS
FT /*tag= a
FT /product= "Serine protease inhibitor"
FT sig_peptide 19..75

```

```

mat_peptide /*tag= b
76..246
/*tag= c
/label= Mature_serine_protease_inhibitor

WO9940183-A1.
XX
XX 12-AUG-1999.
XX
XX 04-FEB-1999; 99WO-US02292.
XX
XX 06-FEB-1998; 98US-0073961.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Ruben SM;
XX
XX WPI; 1999-508502/42.
XX P-PSDB; AAY28645.
XX

New isolated human serine protease and serpin polypeptides, used to
develop products for treating e.g. immune disorders, cancers,
inflammation, transplant rejection or infections, or as food
additives

Claim 4; Page 88; 99pp; English.

The present sequence is a human cdna clone HKABR62 which encodes serine
protease inhibitor (serpin) and is obtained from human keratinocyte
tissue cdna library. The protein obtained from this clone shows a high
degree of sequence similarity to Mus musculus secretory leukocyte
protease inhibitor. The cdna and the serpin are used in the diagnosis
and treatment of disorders related to abnormal level of the protein or
mutation in the nucleotide sequence. The serpin can be used for treating
disorders characterised by degradation of extracellular matrix, e.g.
cancer, arthritis, multiple sclerosis and immune system disorders, for
treating wasting associated with excessive protease production during
inflammation or neurodegenerative disorders e.g. Kallmann's and Down's
syndromes, Alzheimer's and Huntington's diseases. It may also be used
to reduce excess levels of prolactin in the treatment of galactorrhea and
hypogonadism, and decrease the amount of free circulating somatostatin to
prevent somatostatin's inhibitory effect on the release of growth
hormone. The fusion of this protein to His-tag, HA-tag, Igg domains,
etc. facilitates protein purification and fusion to Igg-1, Igg-3 and
albumin increases the half life time in vivo.

SQ Sequence 478 BP; 115 A; 121 C; 135 G; 107 T; 0 other;

Alignment Scores:
Pred. No.: 0.000459 Length: 478
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 20 Gaps: 2

09-833799-13B (1-57) x AAx80909 (1-478)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 76 GCTGTGGAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAAC 132
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 133 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGCAGGACTCTGTGGGG 189
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 190 GAAAGGAGAGTGTGTTACCTGCTGCTGCTTCAAGTGTGTTGTTCT 237

RESULT 67
AAA99904
ID AAA99904 standard; cdna; 762 BP.

```

```

XX AC AAA99904;
XX DT 26-JAN-2001 (first entry)
XX DE cDNA encoding human protein PRO844.
XX KW Cardiovascular; endothelial; angiogenic disorder; PRO179;
XX KW PRO238; PRO364; PRO844; PRO846; PRO1760; PRO205; PRO321; PRO333;
XX KW PRO840; PRO877; PRO878; PRO879; PRO882; PRO885; PRO887;
XX KW gene therapy; ss.
XX OS Homo sapiens.
XX FH Key
XX FT CDS 5..339
XX FT Location/Qualifiers
XX FT /*tag= a
XX PN WO200053757-A2.
XX PD 14-SEP-2000.
XX PF 24-FEB-2000; 2000WO-US05004.
XX PR 08-MAR-1999; 99WO-US05028.
XX PR 12-MAR-1999; 99US-0123957.
XX PR 02-JUN-1999; 99WO-US12252.
XX PR 20-JUL-1999; 99US-0144758.
XX PR 26-JUL-1999; 99US-0145698.
XX PR 01-SEP-1999; 99WO-US20111.
XX PR 15-SEP-1999; 99WO-US21090.
XX PR 30-NOV-1999; 99WO-US28313.
XX PR 30-NOV-1999; 99WO-US28409.
XX PR 02-DEC-1999; 99WO-US28565.
XX PR 18-FEB-2000; 2000WO-US04342.
XX PR 22-FEB-2000; 2000WO-US04414.
XX FA (GETH ) GENENTECH INC.
XX PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Gerriksen ME;
XX PI Goddard A, Gurney AL, Hillan KJ, Marsters SA, Paoni NF, Pitti RM;
XX PI Watanabe CK, Williams PM, Wood WI;
XX DR WPI; 2000-611444/58.
XX DR P-PSDB; AAB27652.
XX PT Novel PRO polypeptides and agonists and antagonists of them, used to
XX PT diagnose and treat cardiovascular, endothelial and angiogenic disorders
XX PT -
XX PS Claim 60; Fig 7; 181pp; English.
XX CC The present invention relates to methods for stimulating or inhibiting
XX CC angiogenesis and cardiovascularization. The methods involve the use of
XX CC pharmaceutical compositions based on the following proteins, PRO179,
XX CC PRO238, PRO364, PRO844, PRO846, PRO1760, PRO205, PRO321, PRO333,
XX CC PRO840, PRO877, PRO878, PRO879, PRO882, PRO885 or PRO887. These
XX CC proteins were identified by isolating cDNA clones encoding secreted
XX CC proteins. The proteins of the invention may be used to diagnose and
XX CC treat cardiovascular, endothelial or angiogenic disorders. The present
XX CC sequence is a cDNA clone encoding one of the proteins of the invention.
XX SQ Sequence 762 BP; 197 A; 194 C; 194 G; 177 T; 0 other;

Alignment Scores:
Pred. No.: 0.000813 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.5% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 21 Gaps: 2

09-833799-13B (1-57) x AAA99904 (1-762)

```

```

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
DB 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCCTCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
DB 119 GTACCTCTCTCAAGTCGATCCTCC---CAGTGTACACAGACACAGGACTGCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
DB 176 GAAAGGAAGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCTCT 223

RESULT 68
AAZ65078
ID AAZ65078 standard; cDNA; 762 BP.
XX AAZ65078;
XX AC AAZ65078;
XX DT 05-APR-2000 (first entry)
XX DE Membrane-bound protein PRO844 encoding cDNA.
XX KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
XX KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX OS Homo sapiens.
XX PN WO9963088-A2.
XX PD 09-DEC-1999.
XX PF 02-JUN-1999; 99WO-US12252.
XX PR 02-JUN-1998; 98US-0087607.
XX PR 02-JUN-1998; 98US-0087759.
XX PR 03-JUN-1998; 98US-0087827.
XX PR 04-JUN-1998; 98US-0088021.
XX PR 04-JUN-1998; 98US-0088025.
XX PR 04-JUN-1998; 98US-0088028.
XX PR 04-JUN-1998; 98US-0088029.
XX PR 04-JUN-1998; 98US-0088030.
XX PR 04-JUN-1998; 98US-0088033.
XX PR 04-JUN-1998; 98US-0088326.
XX PR 05-JUN-1998; 98US-0088167.
XX PR 05-JUN-1998; 98US-0088202.
XX PR 05-JUN-1998; 98US-0088212.
XX PR 05-JUN-1998; 98US-0088217.
XX PR 09-JUN-1998; 98US-0088655.
XX PR 10-JUN-1998; 98US-0088722.
XX PR 10-JUN-1998; 98US-0088730.
XX PR 10-JUN-1998; 98US-0088734.
XX PR 10-JUN-1998; 98US-0088738.
XX PR 10-JUN-1998; 98US-0088740.
XX PR 10-JUN-1998; 98US-0088741.
XX PR 10-JUN-1998; 98US-0088742.
XX PR 10-JUN-1998; 98US-0088810.
XX PR 10-JUN-1998; 98US-0088811.
XX PR 10-JUN-1998; 98US-0088824.
XX PR 10-JUN-1998; 98US-0088825.
XX PR 10-JUN-1998; 98US-0088826.
XX PR 11-JUN-1998; 98US-0088858.
XX PR 11-JUN-1998; 98US-0088861.
XX PR 11-JUN-1998; 98US-0088863.
XX PR 11-JUN-1998; 98US-0088876.
XX PR 12-JUN-1998; 98US-0089090.
XX PR 12-JUN-1998; 98US-0089105.
XX PR 16-JUN-1998; 98US-0089440.
XX PR 16-JUN-1998; 98US-0089512.
XX PR 16-JUN-1998; 98US-0089514.
XX PR 17-JUN-1998; 98US-0089532.
XX PR 17-JUN-1998; 98US-0089538.

```


CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression
 CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.

XX
 SQ Sequence 411 BP; 98 A; 97 C; 94 G; 94 T; 28 other;

Alignment Scores:
 Pred. No.: 0.00144 Length: 411
 Score: 107.00 Matches: 21
 Percent Similarity: 48.00% Conservative: 3
 Best Local Similarity: 42.00% Mismatches: 25
 Query Match: 32.72% Indels: 1
 DB: 24 Gaps: 0

09-833799-13B (1-57) x ABK64114 (1-411)

QY 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27

DB 341 CCACACAGGNGNACCCCTCGNAGTNCNAGTGANTTTGGCCATGTTCGATCTNAA 282

QY 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluG1 47

DB 281 CCCCCCAATTTCTGTGAGATGATGCGCCAGTGCAGCTGACTTCAAGTGTTCATGGG 222

QY 47 YSerCysGlyMetAlaCysPheValPro 56

DB 221 CATGTGTGGGAAATCTCGCTTTCCCT 194

RESULT 71

ID ABL66677/c

XX ABL66677 standard; DNA; 411 BP.

AC ABL66677;

DT 15-MAY-2002 (first entry)

DE Lung cancer related gene sequence SEQ ID NO:5014.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.

OS Homo sapiens.

XX WO200194629-A2.

PN 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 28-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a
 chemical agent to be tested for anti-neoplastic activity, and
 determining a change in expression of a gene of a signature gene set

Claim 1; SEQ ID 5014; 44pp; English.

The present invention describes a method (M1) for screening for an
 anti-neoplastic agent. The method involves exposing cells to a chemical
 agent to be tested for anti-neoplastic activity, determining a change in
 expression of at least one gene (I) of a signature gene set, where (I)
 comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 to ABL70110), or is at least 95% identical to (S), where a change in
 expression is indicative of anti-neoplastic activity. (I) has cytostatic
 activity and can be used in gene therapy. M1 can be used for screening
 an anti-neoplastic agent, and can be used for producing a product which
 is the data collected with respect to the anti-neoplastic agent as a
 result of M1, and the data is sufficient to convey the chemical
 structure and/or properties of the agent. M1 can be used in the
 treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

SQ Sequence 411 BP; 98 A; 97 C; 94 G; 94 T; 28 other;

Alignment Scores:

Pred. No.: 0.00144

Length: 411

Score: 107.00

Matches: 21

```

Percent Similarity: 48.00%      Conservative: 3
Best Local Similarity: 42.00%    Mismatches: 25
Query Match: 32.72%             Indels: 1
DB: 24                           Gaps: 0

09-833799-13B (1-57) x ABL66677 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAAGGNGNAACCCCTGGNAAGTNCNAGTGANTTGGCCAATGTTNGATGCTNAA 282
Qy 27 nProProAsnArgCysLeuLysAspThrAspProGlyIleLysLysCysGluG1 47
Db 281 CCCCCCAATTCTGTGTGAGATGGAGCGGCAAGCGTGAAGTGAAGTGTTCATGGG 222
Qy 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGGAATCTGGGTTCCCT 194

RESULT 72
ABL67461/c
ID ABL67461 standard; DNA; 411 BP.
XX AC ABL67461;
XX DT 15-MAY-2002 (first entry)
XX DE Thyroid cancer related gene sequence SEQ ID NO:5798.
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX KW gene; ds.
XX OS Homo sapiens.
XX FN WO200194629-A2.
XX PD 13-DEC-2001.
XX PF 30-MAY-2001; 2001WO-US10838.
XX PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.

02-OCT-2000; 2000US-237278P.
02-OCT-2000; 2000US-237294P.
02-OCT-2000; 2000US-237295P.
02-OCT-2000; 2000US-237316P.
03-OCT-2000; 2000US-237425P.
03-OCT-2000; 2000US-237598P.
03-OCT-2000; 2000US-237604P.
03-OCT-2000; 2000US-237606P.
03-OCT-2000; 2000US-237608P.
01-NOV-2000; 2000US-244867P.
01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX PA
XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX XX
DR WPI; 2002-188264/24.
XX XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX XX
PS Claim 1; SEQ ID 5798; 44pp; English.
XX XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX XX
SQ Sequence 411 BP; 98 A; 97 C; 94 G; 94 T; 28 other;

Alignment Scores:
Pred. No.: 0.00144      Length: 411
Score: 107.00           Matches: 21
Percent Similarity: 48.00%      Conservative: 3
Best Local Similarity: 42.00%    Mismatches: 25
Query Match: 32.72%             Indels: 1
DB: 24                     Gaps: 0

09-833799-13B (1-57) x ABL67461 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAAGGNGNAACCCCTGGNAAGTNCNAGTGANTTGGCCAATGTTNGATGCTNAA 282
Qy 27 nProProAsnArgCysLeuLysAspThrAspProGlyIleLysLysCysGluG1 47
Db 281 CCCCCCAATTCTGTGTGAGATGGAGCGGCAAGCGTGAAGTGAAGTGTTCATGGG 222
Qy 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGGAATCTGGGTTCCCT 194

RESULT 73
ABL67811/c
ID ABL67811 standard; DNA; 411 BP.
XX AC ABL67811;
XX XX

```

DT	15-MAY-2002	(first entry)	
XX	Ovary cancer related gene sequence SEQ ID NO:6148.		
DE	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;		
XX	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cystosatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;		
KW	gene; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200194629-A2.		
XX			
PD	13-DEC-2001.		
XX			
PF	30-MAY-2001; 2001WO-US10838.		
XX			
PR	05-JUN-2000; 2000US-209473P.		
PR	05-JUN-2000; 2000US-209531P.		
PR	18-SEP-2000; 2000US-231133P.		
PR	18-SEP-2000; 2000US-233617P.		
PR	20-SEP-2000; 2000US-234009P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	20-SEP-2000; 2000US-234052P.		
PR	22-SEP-2000; 2000US-234059P.		
PR	22-SEP-2000; 2000US-234567P.		
PR	25-SEP-2000; 2000US-234923P.		
PR	25-SEP-2000; 2000US-234924P.		
PR	25-SEP-2000; 2000US-235077P.		
PR	25-SEP-2000; 2000US-235082P.		
PR	25-SEP-2000; 2000US-235134P.		
PR	25-SEP-2000; 2000US-235280P.		
PR	26-SEP-2000; 2000US-235637P.		
PR	26-SEP-2000; 2000US-235638P.		
PR	27-SEP-2000; 2000US-235711P.		
PR	27-SEP-2000; 2000US-235720P.		
PR	27-SEP-2000; 2000US-235840P.		
PR	27-SEP-2000; 2000US-235863P.		
PR	28-SEP-2000; 2000US-236028P.		
PR	28-SEP-2000; 2000US-236032P.		
PR	28-SEP-2000; 2000US-236033P.		
PR	28-SEP-2000; 2000US-236034P.		
PR	28-SEP-2000; 2000US-236109P.		
PR	28-SEP-2000; 2000US-236111P.		
PR	29-SEP-2000; 2000US-236842P.		
PR	29-SEP-2000; 2000US-236891P.		
PR	02-OCT-2000; 2000US-237172P.		
PR	02-OCT-2000; 2000US-237173P.		
PR	02-OCT-2000; 2000US-237278P.		
PR	02-OCT-2000; 2000US-237294P.		
PR	02-OCT-2000; 2000US-237295P.		
PR	02-OCT-2000; 2000US-237316P.		
PR	03-OCT-2000; 2000US-237425P.		
PR	03-OCT-2000; 2000US-237598P.		
PR	03-OCT-2000; 2000US-237604P.		
PR	03-OCT-2000; 2000US-237606P.		
PR	03-OCT-2000; 2000US-237608P.		
PR	01-NOV-2000; 2000US-244867P.		
PR	01-NOV-2000; 2000US-245084P.		
XX			
PA	(AVAL-) AVALON PHARM.		
XX			
PI	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
XX			
DR	WPI; 2002-188264/24.		
XX			
PT	Screening for anti-neoplastic agent involves exposing cells to a		
PT	chemical agent to be tested for anti-neoplastic activity, and		
PT	determining a change in expression of a gene of a signature gene set		
XX			
PS	Claim 1; SEQ ID 6148; 44pp; English.		
XX			

XX Immunogenic portion of an ovarian carcinoma protein and the nucleic
PT acid encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer -
XX Claim 18; Page 192; 299pp; English.
XX The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines.
CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
CC are useful for the prevention, diagnosis and treatment of cancer,
CC preferably ovarian cancer. AAB69691 to AAA70077 and AAB12552 to
CC represent human ovarian carcinoma polynucleotides and proteins used in
CC the exemplification of the present invention.
XX SQ Sequence 292 BP; 63 A; 71 C; 75 G; 83 T; 0 other;
Alignment Scores:
Pred. No.: 0.00161 Length: 292
Score: 105.00 Matches: 18
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 51.43% Mismatches: 15
Query Match: 32.11% Indels: 0
DB: 21 Gaps: 0
09-833799-13B (1-57) x AAA70066 (1-292)
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1 CAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGCGTGAC 60
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 61 TTGAAGTGTTCATGGCGATGTGGGAAATCCTGCGTTTCCCT 105
RESULT 75
ABN72960
ID ABN72960 standard; DNA; 292 BP.
XX AC ABN72960;
XX 02-JUL-2002 (first entry)
XX Ovarian carcinoma antigen polynucleotide #65.
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX Homo sapiens.
XX WO200206317-A2.
XX 24-JAN-2002.
XX 17-JUL-2001; 2001WO-US22635.
XX 17-JUL-2000; 2000US-0617747.
XX 10-AUG-2000; 2000US-0636801.
XX 20-SEP-2000; 2000US-0667857.
XX 04-APR-2001; 2001US-0827271.
XX 18-JUN-2001; 2001US-0884441.
XX (CORI-) CORIXA CORP.
XX Mittham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedwick TS, Carter D, Hill P, Albone E;
XX WPI; 2002-164781/21.
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer -

XX Example 2; Page 310; 408pp; English.
XX This invention relates to polypeptides comprising an immunogenic
CC portion of an ovarian carcinoma protein which acts as an
CC immunostimulant and is cytostatic. The polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells that express the polypeptides are useful for stimulating an
CC immune response in a patient and treating ovarian cancer. This
CC sequence represents DNA related to the invention.
XX SQ Sequence 292 BP; 63 A; 71 C; 75 G; 83 T; 0 other;
Alignment Scores:
Pred. No.: 0.00161 Length: 292
Score: 105.00 Matches: 18
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 51.43% Mismatches: 15
Query Match: 32.11% Indels: 0
DB: 24 Gaps: 0
09-833799-13B (1-57) x ABN72960 (1-292)
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1 CAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGCGTGAC 60
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 61 TTGAAGTGTTCATGGCGATGTGGGAAATCCTGCGTTTCCCT 105
Search completed: February 15, 2003, 20:26:35
Job time : 232 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 20:11:56 ; Search time 1458 Seconds
(without alignments)
633.157 Million cell updates/sec

Title: 09-833799-13B

Perfect score: 327

Sequence: 1 aqepvgpvtkpgscpill.....cpgikkccgscgmcfvpq 57

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09833799/runat_12022003_120042_23664/app_query.fasta_1.199
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799 @CGN 1 1 2874 @runat_12022003_120042_23664 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: gb_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
1	327	100.0	219	10	BE182652	BE182652 RC3-HT064
2	327	100.0	289	10	BE181373	BE181373 CM0-HT063
3	327	100.0	312	10	BE181959	BE181959 CM1-HT064
4	327	100.0	340	13	EG987761	EG987761 CM3-HT119
5	327	100.0	342	10	BE182993	BE182993 CM4-HT063
6	327	100.0	344	10	BE181546	BE181546 CM4-HT063
7	327	100.0	359	10	BE182864	BE182864 CM4-HT065
8	327	100.0	364	12	BE175602	BE175602 CM4-HT074
9	327	100.0	370	12	BE175605	BE175605 CM4-HT074
10	327	100.0	370	12	BE175626	BE175626 CM4-HT074
11	327	100.0	370	12	BE172160	BE172160 CM4-HT010
12	327	100.0	371	12	BF838555	BF838555 IL5-HT099
13	327	100.0	371	12	BE772166	BE772166 CM4-F010
14	327	100.0	375	9	A1285400	A1285400 qt69f06.x
15	327	100.0	379	12	BE175603	BE175603 CM4-HT074
16	327	100.0	380	12	BE772155	BE772155 CM4-F010
17	327	100.0	380	12	BE772171	BE772171 CM4-F010
18	327	100.0	386	10	BE182879	BE182879 CM4-HT065
19	327	100.0	386	12	BE772169	BE772169 CM4-F010
20	327	100.0	388	9	AA583997	AA583997 nn65a01.s
21	327	100.0	388	12	BE772159	BE772159 CM4-F010
22	327	100.0	389	12	BE772168	BE772168 CM4-F010
23	327	100.0	390	9	A1393896	A1393896 tg05e09.x
24	327	100.0	390	9	A1989737	A1989737 wt21e08.x
25	327	100.0	390	12	BE175611	BE175611 CM4-HT074
26	327	100.0	395	12	BE175615	BE175615 CM4-HT074
27	327	100.0	397	12	BE772137	BE772137 CM4-F010
28	327	100.0	399	9	AA586974	AA586974 nn69d08.s
29	327	100.0	400	12	BE175592	BE175592 CM4-HT074
30	327	100.0	403	12	BE175609	BE175609 CM4-HT074
31	327	100.0	405	9	A1874186	A1874186 wt50a02.x
32	327	100.0	406	10	AW137392	AW137392 UI-H-B11
33	327	100.0	406	12	BE175604	BE175604 CM4-HT074
34	327	100.0	418	9	A1392753	A1392753 tg23g03.x
35	327	100.0	419	9	AA586943	AA586943 nn69a01.s
36	327	100.0	422	10	AW082097	AW082097 xb60d08.x
37	327	100.0	428	12	BE772154	BE772154 CM4-F010
38	327	100.0	429	9	A1283910	A1283910 qt66a10.x
39	327	100.0	437	12	BF836224	BF836224 QV3-HT101
40	327	100.0	438	10	AW001880	AW001880 wt81d06.x
41	327	100.0	438	12	BE175601	BE175601 CM4-HT074
42	327	100.0	441	12	BF836743	BF836743 CM2-HT096
43	327	100.0	448	9	A1582329	A1582329 tq66g09.x
44	327	100.0	453	9	A1459240	A1459240 tk11c09.x
45	327	100.0	465	12	BF824934	BF824934 ILO-HN003
46	327	100.0	466	12	BF824930	BF824930 ILO-HN003
47	327	100.0	466	12	BF824935	BF824935 ILO-HN003
48	327	100.0	467	12	BF824920	BF824920 ILO-HN003
49	327	100.0	467	12	BF825207	BF825207 ILO-HN003
50	327	100.0	471	12	BF837734	BF837734 QV3-HT101
51	327	100.0	477	12	BF824932	BF824932 ILO-HN003
52	327	100.0	480	9	AA582866	AA582866 nn72a09.s
53	327	100.0	487	12	BF825204	BF825204 ILO-HN003
54	327	100.0	490	12	BF002099	BF002099 7G99b04.x
55	327	100.0	498	12	BF837735	BF837735 QV3-HT101
56	327	100.0	522	14	BQ941085	BQ941085 AGENCOURT
57	327	100.0	565	12	BG490472	BG490472 602519610
58	324	99.1	321	12	BE772161	BE772161 CM4-F010
59	324	99.1	421	13	BI061069	BI061069 IL3-UT011
60	323	98.8	263	12	BE711100	BE711100 RC3-HT064
61	323	98.8	420	9	A1924155	A1924155 wn64c08.x
62	323	98.8	447	12	BF835679	BF835679 QV3-HT101
63	323	98.8	453	10	AW845140	AW845140 ILO-CT000
64	323	98.8	455	10	AW845142	AW845142 ILO-CT000

```

c 65 322 98.5 444 9 A1858070
c 66 321 98.2 342 12 BE715619
c 67 321 98.2 374 12 BF836201
c 68 321 98.2 501 12 BF825196
c 69 320 97.9 316 12 BE711083
c 70 320 97.9 344 12 AW844969
c 71 320 97.9 428 14 BM768573
c 72 319 97.6 394 13 BI021805
c 73 318 97.2 386 9 A586983
c 74 318 97.2 419 12 BF736734
c 75 318 97.2 423 9 A1830872
c 76 318 97.2 527 9 A1807596
c 77 316 96.6 431 12 BF837730
c 78 315 96.3 298 12 BF837690
c 79 315 96.3 313 12 BF736294
c 80 315 96.3 314 12 BF837677
c 81 315 96.3 345 12 BF094327
c 82 314 96.0 275 12 BE711090
c 83 314 96.0 391 12 BE772162
c 84 314 96.0 393 10 BE183902
c 85 314 96.0 448 12 BE772164
c 86 314 96.0 430 12 AW082007
c 87 313 95.7 467 12 BF836725
c 88 311 95.1 289 12 BF834601
c 89 309 94.5 275 10 BE182650
c 90 308 94.2 344 10 AW860590
c 91 307 93.9 445 9 A1830232
c 92 305 93.3 470 12 BF913869
c 93 302 92.4 400 9 A1833796
c 94 300 91.7 555 9 A583567
c 95 298 91.1 411 9 A586718
c 96 298 91.1 478 9 A527557
c 97 291 89.0 190 10 BE181433
c 98 290.5 88.8 468 12 BF090376
c 99 286 87.5 299 10 BE184317
100 284 86.9 470 12 BF913875

FEATURES
source
1..219
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 51 a 60 c 59 g 49 t
ORIGIN
1..44e-29 Length: 219
Pred. No.: 327.00 Matches: 57
Score: 327.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 10

09-833799-13B (1-57) x BE182652 (1-219)
QY 1 AlaglnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 24 GCGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCCTGGCTGCTCCCAATTATCTTG 83
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 84 ATCCGTGGCGCCATGTGAATCCCCCTAACCCGTGCTGAAAGATACCTGACTGCCCCAGGA 143
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 144 ATCAAGAGTGTCTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCCAG 194

RESULT 2
BE181373/c
LOCUS BE181373 289 bp mRNA linear EST 22-JUN-2000
DEFINITION CM0-HT0633-240300-304-g12 HT0633 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE181373
VERSION BE181373.1 GI:8660549
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 219)
REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC3-HT0649-090

TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC3-HT0649-090

```

Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=CM0-HT0633-240
 300-304-g12&t3=2000-03-24&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 17
 High quality sequence stop: 289.
 Location/Qualifiers

FEATURES

source
 1..289
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0633"
 /dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 62 a 74 c 93 g 60 t

ORIGIN

Alignment Scores:
 Pred. No.: 1.99e-29 Length: 289
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

09-833799-13B (1-57) x BE181373 (1-289)

Qy 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 274 GCGCAGGAGCAGTCAAGGTCAGTCTCCACTAAGCTCGCTCCGCCCATATCTTG 215
 Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 214 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGACTGCCCAGGA 155
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 154 ATCAAGAAGTGTCTGAAGGCTCTTGGGGATGGCCCTGTTTCGTTCCCCAG 104

RESULT 3

BE181959 312 bp mRNA linear EST 22-JUN-2000
 LOCUS CMI-HT0641-280400-206-d02 HT0641 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE181959
 ACCESSION BE181959.1 GI:8661135
 VERSION EST.
 KEYWORDS human.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=CM1-HT0641-280
 400-206-d02&t3=2000-04-28&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 15
 High quality sequence stop: 312.
 Location/Qualifiers

FEATURES

source
 1..312
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0641"
 /dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 70 a 93 c 81 g 68 t

ORIGIN

Alignment Scores:
 Pred. No.: 2.18e-29 Length: 312
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

09-833799-13B (1-57) x BE181959 (1-312)

Qy 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 71 GCGCAGGAGCAGTCAAGGTCAGTCTCCACTAAGCTCGCTCCGCCCATATCTTG 130
 Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 131 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGACTGCCCAGGA 190
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 191 ATCAAGAAGTGTCTGAAGGCTCTTGGGGATGGCCCTGTTTCGTTCCCCAG 241

RESULT 4

BE181959 340 bp mRNA linear EST 13-JUN-2001
 LOCUS CMI-HT1190-110101-624-h04 HT1190 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE181959
 ACCESSION BE181959.1 GI:14391831
 VERSION EST.
 KEYWORDS human.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-HT1190-110101-624-h04&t3=2001-01-11&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 6
 High quality sequence stop: 340.
 Location/Qualifiers

FEATURES

source

1. .340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT1190"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 79 a 82 c 97 g 82 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2 41e-29 Length: 340
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

09-833799-13B (1-57) x BG987761 (1-340)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 238 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCATTTATCTTG 179
 |||||
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 178 ATCCGGTGGCCCATGTGTAATCCCCCTAACCCGCTGCTTGAAGATACTGACTGCCCGAGA 119
 |||||
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 118 ATCAAGAAGTGCTGTGAAGCTCTTGGGATGGCTGTTTCGTTCCCGAG 68
 |||||

RESULT 5

BE182993/c

LOCUS CM4-HT0653-180400-146-all HT0653 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE182993
 ACCESSION BE182993.1 GI:8662169
 VERSION
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

20202663

COMMENT

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4-HT0653-180400-146-all&t3=2000-04-18&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 342.
 Location/Qualifiers

FEATURES

source

1. .342
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0653"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 80 a 82 c 100 g 79 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2 43e-29 Length: 342
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

09-833799-13B (1-57) x BE182993 (1-342)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 259 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCATTTATCTTG 200
 |||||
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 199 ATCCGGTGGCCCATGTGTAATCCCCCTAACCCGCTGCTTGAAGATACTGACTGCCCGAGA 140
 |||||
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 139 ATCAAGAAGTGCTGTGAAGCTCTTGGGATGGCTGTTTCGTTCCCGAG 89
 |||||

RESULT 6

BE181546/c

LOCUS CM0-HT0634-270300-308-f02 HT0634 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE181546
 ACCESSION BE181546.1 GI:8660722
 VERSION
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0634-270
300-308-f02&t3=2000-03-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 344.
Location/Qualifiers

FEATURES

source
1..344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0634"
/dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 79 a 80 c 113 g 72 t
ORIGIN

Alignment Scores:
Pred. No.: 2,45e-29 Length: 344
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x BE181546 (1-344)

Qy 1 AlAGlnGluProValysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 294 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTCGCTGCCCTATATCTTG 235
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 234 ATCCGGTGGCCATGTTGAATCCCTTAAACGGCTGTTGAAGATGACTGATGCCAGGA 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 174 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTGTTTGGTTCCCCAG 124

RESULT 7
BE182864
LOCUS BE182864 CM4-HT0652-150400-143-a07 HT0652 Homo sapiens cdna, mRNA sequence.
ACCESSION BE182864.1 GI:8662040
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

TITLE
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

COMMENT

20202663

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0652-150
400-143-a07&t3=2000-04-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 359.
Location/Qualifiers

1..359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 92 a 90 c 95 g 82 t
ORIGIN

FEATURES

source

1..359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 92 a 90 c 95 g 82 t
ORIGIN

Alignment Scores:
Pred. No.: 2,57e-29 Length: 359
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x BE182864 (1-359)

Qy 1 AlAGlnGluProValysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 172 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTCGCTGCCCTATATCTTG 231
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 232 ATCCGGTGGCCATGTTGAATCCCTTAAACGGCTGTTGAAGATGACTGATGCCAGGA 291
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 292 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTGTTTGGTTCCCCAG 342

RESULT 8

BE1715602/c

LOCUS BE1715602 CM4-HT0744-150600-201-a06 HT0744 Homo sapiens cdna, mRNA sequence.
ACCESSION BE1715602
VERSION BE1715602.1 GI:10103867
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 364)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

TITLE

sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE
 COMMENT

contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
 600-201-a06&t3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 20
 High quality sequence stop: 364.

FEATURES
 source
 1..364
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 85 a 92 c 100 g 87 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.61e-29 Length: 364
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715602 (1-364)

QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 239 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTGCTGCCCATTAATCTTG 180
 |||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||

Db 179 ATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTTGAAAGATACACTGACTGCCCAGGA 120
 |||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||

Db 119 ATCAAGAGTGTGTGAAGGCTTTGCGGGATGGCTGTTTCGTTCCCCAG 69
 |||||

RESULT 9
 BE715605/c 370 bp mRNA linear EST 12-SEP-2000
 LOCUS
 CM4-HT0744-160600-201-a10 HT0744 Homo sapiens cDNA, mRNA sequence.
 BE715605
 ACCESSION
 BE715605.1 GI:10103870
 VERSION
 EST.
 BE715605.1
 KEYWORDS
 human.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 370)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldmann,G.H., Carvalhal,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE
 COMMENT

contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
 600-201-a10&t3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 11
 High quality sequence stop: 370.

FEATURES
 source
 1..370
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 85 a 97 c 99 g 89 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.66e-29 Length: 370
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715605 (1-370)

QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||

Db 240 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTGCTGCCCATTAATCTTG 181
 |||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||

Db 180 ATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTTGAAAGATACACTGACTGCCCAGGA 121
 |||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||

Db 120 ATCAAGAGTGTGTGAAGGCTTTGCGGGATGGCTGTTTCGTTCCCCAG 70
 |||||

RESULT 10
 BE715626 370 bp mRNA linear EST 12-SEP-2000
 LOCUS
 CM4-HT0744-160600-201-h09 HT0744 Homo sapiens cDNA, mRNA sequence.
 BE715626
 ACCESSION
 BE715626.1 GI:10103891
 VERSION
 EST.
 BE715626.1
 KEYWORDS
 human.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 370)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldmann,G.H., Carvalhal,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
600-201-h09&t3=2000-06-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 370.

FEATURES

source
1..370
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 85 a 110 c 87 g 88 t
ORIGIN

Alignment Scores:
Pred. No.: 2.66e-29 Length: 370
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715626 (1-370)

Qy 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 97 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCTGCCCATATCTTG 156
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIlyAspThrAspCysProGly 40
Db 157 ATCCGGTGGCCAGTGTGAATCCCCCTTAACCGCTGCTTGAAGAGATCTGACTGCCACGGA 216
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 217 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGGCCCTGTTCTGCCCCAG 267

RESULT 11
BE772160/c 370 bp mRNA linear EST 20-SEP-2000
LOCUS BE772160
DEFINITION CM4-FT0104-230600-215-c08 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772160
VERSION BE772160.1 GI:10225818
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE

JOURNAL
MEDLINE
COMMENT

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-c08&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 370.

FEATURES

source
1..370
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 95 c 99 g 90 t
ORIGIN

Alignment Scores:
Pred. No.: 2.66e-29 Length: 370
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772160 (1-370)

Qy 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 239 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCTGCCCATATCTTG 180
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIlyAspThrAspCysProGly 40
Db 179 ATCCGGTGGCCAGTGTGAATCCCCCTTAACCGCTGCTTGAAGAGATCTGACTGCCACGGA 120
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 119 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGGCCCTGTTCTGCCCCAG 69

RESULT 12
BF838555 371 bp mRNA linear EST 13-JAN-2001
LOCUS BF838555
DEFINITION IL5-HT0990-251100-263-g08 HT0990 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF838555
VERSION BF838555.1 GI:12190853
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 371)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldmann,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-HT0990-251100-263-g08&t3=2000-11-25&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 367.
 FEATURES
 source
 1..371
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0990"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 82 a 117 c 85 g 87 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.67e-29 Length: 371
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 09-833799-13B (1-57) x BF838555 (1-371)
 QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 77 GCGCAAGACCCAGTCAAGGTCAGCTCCACTAAGCTGGCTCTGCCCATTTATCTTG 136
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 137 ATCCGGTGCAGCATGTGAATCCCCCTAACCGCTCTTGAAGATAGTACTGACGCCAGGA 196
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 197 ATCAAGAAAGTGTGAAGGCTCTTTCGGGGATGGCTGTTCGTTCCCCAG 247
 RESULT 13
 BE772166/c BE772166/c 371 bp mRNA linear EST 20-SEP-2000
 LOCUS CM4-FT0104-230600-215-d05 FT0104 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE772166
 ACCESSION BE772166
 VERSION BE772166.1 GI:10225824
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 371)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldmann,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-HT0990-251100-263-g08&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 371.
 FEATURES
 source
 1..371
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 84 a 96 c 101 g 90 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.67e-29 Length: 371
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 09-833799-13B (1-57) x BE772166 (1-371)
 QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 240 GCGCAAGACCCAGTCAAGGTCAGCTCCACTAAGCTGGCTCTGCCCATTTATCTTG 181
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 180 ATCCGGTGCAGCATGTGAATCCCCCTAACCGCTCTTGAAGATAGTACTGACGCCAGGA 121
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 120 ATCAAGAAAGTGTGAAGGCTCTTTCGGGGATGGCTGTTCGTTCCCCAG 70
 RESULT 14
 AI285400/c AI285400/c 375 bp mRNA linear EST 23-NOV-1998
 LOCUS qt69f06.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:1960547.3,
 DEFINITION similar to gb:SS8717 rna1 ELAFIN PRECURSOR (HUMAN); contains element
 THR repetitive element ;, mRNA sequence.
 ACCESSION AI285400
 VERSION AI285400.1 GI:3923633

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 375)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 356.

FEATURES

source
 1..375
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1960547"
 /clone_lib="NCI CGAP Eso2"
 /tissue_type="squamous cell carcinoma"
 /lab_host="DH10B"
 /note="Organ: esophagus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.1 kb. Life Technologies catalog #11502-010"
BASE COUNT 91 a 83 c 125 g 75 t 1 others
ORIGIN

Alignment Scores:
 Pred. No.: 2,71e-29 Length: 375
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

09-833799-13B (1-57) x A1285400 (1-375)

Qy 1 AlaGlnGluProVallysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 349 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCGCCCATATCTTG 290
 |||||
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnAtgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 289 ATCGGTGCCCATGTGTAATCCCTTACCGCTGCTTGAAGATGACTGACGCCAGGA 230
 |||||
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 229 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCCAG 179
 |||||

RESULT 15
 BE715603/c 379 bp mRNA linear EST 12-SEP-2000
LOCUS CM4-HT0744-160600-201-a07 HT0744 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE715603
ACCESSION BE715603.1 GI:10103868
VERSION EST.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 379)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

TITLE
JOURNAL Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.
MEDLINE Shotgun sequencing of the human transcriptome with ORF expressed
COMMENT sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160>)
 600-201-a07&t3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 25
 High quality sequence stop: 379.

FEATURES

source
 1..379
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 86 a 95 c 112 g 86 t
ORIGIN

Alignment Scores:
 Pred. No.: 2,74e-29 Length: 379
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715603 (1-379)

Qy 1 AlaGlnGluProVallysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 260 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCGCCCATATCTTG 201
 |||||
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnAtgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 200 ATCCGGTGGCCATGTGTAATCCCTTACCGCTGCTTGAAGATGACTGACGCCAGGA 141
 |||||
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 140 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCCAG 90
 |||||

RESULT 16
 BE772155/c 380 bp mRNA linear EST 20-SEP-2000
LOCUS BE772155
DEFINITION CM4-FT0104-230600-215-b05 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772155
VERSION BE772155.1 GI:10225813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 380)

AUTHORS
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., deOliveira,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
 20202663

COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230 600-215-b05&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 365.
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
 89 a 99 c 100 g 92 t

FEATURES
 source
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Alignment Scores:
 Pred. No.: 2.75e-29 Length: 380
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772155 (1-380)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 247 GCGCAAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCATATCTTG 188
 |||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||

Db 187 ATCCGGTGGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAAGATACACTGACGCCGAGGA 128
 |||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||

Db 127 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 77
 |||||

RESULT 17
 BE772171/c
 LOCUS BE772171 380 bp mRNA linear EST 20-SEP-2000
 DEFINITION CM4-FT0104-230600-215-g07 FT0104 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE772171
 VERSION BE772171.1 GI:10225829
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 380)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., deOliveira,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230 600-215-g07&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 23
 High quality sequence stop: 380.
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
 89 a 97 c 106 g 88 t

Alignment Scores:
 Pred. No.: 2.75e-29 Length: 380
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772171 (1-380)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 251 GCGCAAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCATATCTTG 192
 |||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||

Db 191 ATCCGGTGGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAAGATACACTGACGCCGAGGA 132
 |||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||

Db 131 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 81
 |||||

RESULT 18
 BE182879
 LOCUS BE182879 386 bp mRNA linear EST 22-JUN-2000
 DEFINITION CM4-HT0652-150400-143-g08 HT0652 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE182879
 VERSION BE182879.1 GI:8662055
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 386)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0652-150
 400-143-g08&t3=2000-04-15&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 61.
 FEATURES
 source Location/Qualifiers
 1..386
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0652"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site: 1: Smal;
 Site_2: Smal; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 136,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 89 a 104 c 105 g 88 t
 ORIGIN
 Alignment Scores: Length: 386
 Pred. No.: 2.8e-29 Matches: 57
 Score: 327.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 10
 09-833799-13B (1-57) x BE182879 (1-386)
 Qy 1 AlaGlnGluProValysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
 Db 146 GCGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCTGGCTGCTCCCATTAATCTTG 205
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIysAspThrAspCysProGly 40
 Db 206 ATCCGGTGGCGCATGTGAAATCCCCCTAACCGCTGCTTGAAGATACACTGACTGCCAGGA 265
 Qy 41 IleIysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 266 ATCAAGAAGTCTGTGAAGGCTTTTCGGGATGGCCTGTTTCGTTCCCGAG 316
 RESULT 19
 BE772169/c 386 bp mRNA linear EST 20-SEP-2000
 LOCUS BE772169 CM4-F0104-230600-215-f02 FT0104 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE772169
 VERSION BE772169.1 GI:10225827
 KEYWORDS EST.

SOURCE Homo sapiens
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 386)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-F0104-230
 600-215-f02&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 45
 High quality sequence stop: 386.
 FEATURES
 source Location/Qualifiers
 1..386
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site: 1: Smal;
 Site_2: Smal; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 90 a 99 c 107 g 90 t
 ORIGIN
 Alignment Scores: Length: 386
 Pred. No.: 2.8e-29 Matches: 57
 Score: 327.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 12
 09-833799-13B (1-57) x BE772169 (1-386)
 Qy 1 AlaGlnGluProValysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
 Db 255 GCGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCTGGCTGCTCCCATTAATCTTG 196
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIysAspThrAspCysProGly 40
 Db 195 ATCCGGTGGCGCATGTGAAATCCCCCTAACCGCTGCTTGAAGATACACTGACTGCCAGGA 136
 Qy 41 IleIysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 135 ATCAAGAAGTCTGTGAAGGCTTTTCGGGATGGCCTGTTTCGTTCCCGAG 85
 RESULT 20
 AA583997/c 388 bp mRNA linear EST 26-SEP-1997
 LOCUS AA583997 nn65a01.sl NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1088712 3,
 DEFINITION similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AA583997
 VERSION AA583997.1 GI:2368606
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 388)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 480 Std Error: 0.00
 Seq primer: -40m13 fwd. Et from Amersham.
 Location/Qualifiers
 FEATURES
 source
 1..388
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1088712"
 /clone_lib="NCI-CGAP_Lar1"
 /tissue_type="larynx"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: larynx; Vector: Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGCGCAGG
 3' 3' adaptor sequence: 5' CTCAGTGTCTTTTCTTTTCTTTT 3'
 Average insert size: 0.9 kb."
 BASE COUNT 91 a 83 c 124 g 90 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,82e-29 Length: 388
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 09-833799-13B (1-57) x AA583997 (1-388)
 QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 363 GCGCAAGAGCCAGTCAAAAGGTCAGTCCAGTCCACTAAGCTGGCTGCTGCCCATTTATCTTG 304
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 303 ATCCGGTGGCCATGTTGAATCCCTTACCGCTGCTTGAAGATACACTGCTGCCAGGA 244
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 243 ATCAAGAGTGTGTGAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 193
 RESULT 21
 BE772159
 LOCUS CM4-FT0104-230600-215-c06 FT0104 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE772159
 ACCESSION BE772159
 VERSION BE772159.1 GI:10225817
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 388)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=et2=CM4-FT0104-230>)
 600-215-c06ft3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 26
 High quality sequence stop: 388.
 Location/Qualifiers
 FEATURES
 source
 1..388
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 88 a 115 c 93 g 92 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,82e-29 Length: 388
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 09-833799-13B (1-57) x BE772159 (1-388)
 QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 115 GCGCAAGAGCCAGTCAAAAGGTCAGTCCAGTCCACTAAGCTGGCTGCTGCCCATTTATCTTG 174
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 175 ATCCGGTGGCCATGTTGAATCCCTTACCGCTGCTTGAAGATACACTGCTGCCAGGA 234
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 235 ATCAAGAGTGTGTGAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 285
 RESULT 22
 BE772168
 LOCUS CM4-FT0104-230600-215-d07 FT0104 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE772168
 ACCESSION BE772168
 VERSION BE772168.1 GI:10225826
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 389)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, A.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=CM4-FT0104-230
600-215-d07&t3=2000-06-23&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 32

High quality sequence stop: 389.

FEATURES

source

1. 389
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT

ORIGIN

91 a 115 c 92 g 91 t

Alignment Scores:

Pred. No.: 2,83e-29 Length: 389
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772168 (1-389)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20

Db 116 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATATCTTG 175

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 176 ATCCGGTGGCCGATGTGAATCCCGCTTAACCGCTGTGAAGATATCTGATGCCACGGA 235

Qy 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 236 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGGCTTGTTCGTTCCCCAG 286

RESULT 23

AI393896/c

LOCUS

AI393896 390 bp mRNA linear EST 30-MAR-1999

DEFINITION tq05e09.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2107912.3,
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ;, mRNA sequence.

ACCESSION

AI393896

VERSION

AI393896.1 GI:4223443

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 390)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,

M.D., Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 802 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 381.

Location/Qualifiers

1. 390

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:2107912"

/tissue_type="B-Cell, chronic lymphocytic leukemia"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5',

TGTTACCAATCTGAAGTGGAGCGCGCGCTGCTTTTCTTTTCTTTTCTTTTCTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

Library is normalized, and was constructed by Bento

Soares and M. Fatima Bonaldo."

BASE COUNT 95 a 86 c 126 g 83 t

ORIGIN

09-833799-13B (1-57) x AI393896 (1-390)

Alignment Scores:

Pred. No.: 2,84e-29 Length: 390
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20

Db 350 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATATCTTG 291

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 290 ATCCGGTGGCCGATGTGAATCCCGCTTAACCGCTGTGAAGATATCTGATGCCACGGA 231

Qy 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 230 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGGCTTGTTCGTTCCCCAG 180

RESULT 24

AI989737/c

LOCUS

AI989737

DEFINITION IMAGE:2520710.3, similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN)
);, mRNA sequence.

ACCESSION

AI989737

```

VERSION AI989737.1 GI:5836618
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
JOURNAL Unpublished (1997)
CONTACT Contact: Robert Strausberg, Ph.D.
EMAIL Email: cgapbs-remail.nih.gov
THIS CLONE is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..390
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2520710"
/clone_lib="Soares Dieckgraefe_colon_NHCD"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATGTAAGTGGAGCGCGCGCTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 94 a 86 c 126 g 84 t
ORIGIN

Alignment Scores:
Pred. No.: 2,84e-29 Length: 390
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI989737 (1-390)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 354 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTCGCCCATATCTTG 295
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 294 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTTGAAAGATACTGACTGCCAGGA 235
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 234 ATCAAGAAGTGTGTAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 184

RESULT 25
LOCUS BE715611/c 390 bp mRNA linear EST 12-SEP-2000
DEFINITION CM4-HT0744-160600-201-c07 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715611
VERSION BE715611.1 GI:10103876
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 390)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM4-HT0744-160600-201-c07&t3=2000-06-16&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 30

High quality sequence stop: 390.

Location/Qualifiers

1..390

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0744"

/dev_stage="Adult"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 90 a 97 c 111 g 92 t

ORIGIN

Alignment Scores:

Pred. No.: 2,84e-29 Length: 390

Score: 327.00 Matches: 57

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715611 (1-390)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20

Db 259 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTCGCCCATATCTTG 200

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 199 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTTGAAAGATACTGACTGCCAGGA 140

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 139 ATCAAGAAGTGTGTAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 89

RESULT 26

LOCUS BE715615

DEFINITION CM4-HT0744-160600-201-d10 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE715615

VERSION BE715615.1 GI:10103880

KEYWORDS EST.

SOURCE ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 395)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
 600-201-d10&t3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 4
 High quality sequence stop: 395.

FEATURES
 source
 1..395
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 91 a 115 c 96 g 93 t

Alignment Scores:
 Pred. No.: 2,88e-29 Length: 395
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715615 (1-395)
 Qy 1 AlaGlnGluProValSerThryLysProGlySerCysProIleLeu 20
 Db 122 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTCGCTCCGCCCATTAATCTTG 181
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 182 ATCCGGTGGCGCATGTTGAATCCCGCTAACCGCTGCTTGAAGATACACTGACTGCCAGGA 241
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 242 ATCAAGAAGTCTGTGAAGGCTCTTCGGGATGCCCTGTTTCGTTCCTCCAG 292

RESULT 27
 BE772137
 LOCUS BE772137 397 bp mRNA linear EST 20-SEP-2000
 DEFINITION CM4-FT0104-230600-215-162 FT0104 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE772137

VERSION BE772137.1 GI:10225795
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 397)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
 600-215-a12&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 4
 High quality sequence stop: 397.

FEATURES
 source
 1..397
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 90 a 118 c 95 g 94 t

Alignment Scores:
 Pred. No.: 2,9e-29 Length: 397
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772137 (1-397)
 Qy 1 AlaGlnGluProValSerThryLysProGlySerCysProIleLeu 20
 Db 124 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTCGCTCCGCCCATTAATCTTG 183
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 184 ATCCGGTGGCGCATGTTGAATCCCGCTAACCGCTGCTTGAAGATACACTGACTGCCAGGA 243
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 244 ATCAAGAAGTCTGTGAAGGCTCTTCGGGATGCCCTGTTTCGTTCCTCCAG 294

RESULT 28
 AA586974/c
 LOCUS AA586974 399 bp mRNA linear EST 26-SEP-1997

DEFINITION	nm69d08.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089135 3' similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element THR repetitive element ;, mRNA sequence.
ACCESSION	AA586974
VERSION	AA586974.1
KEYWORDS	GI:2397788
ORGANISM	human.
SOURCE	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 399)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Stratagene, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. cDNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/obpr/image/image.html Insert Length: 523 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham.

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QW4-HT0744-160
600-201-c04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 403.

FEATURES
source
1..403
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
Derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 96 a 98 c 113 g 96 t
ORIGIN

Alignment Scores:
Pred. No.: 2,95e-29 Length: 403
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715609 (1-403)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 272 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCTATATCTTG 213
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 212 ATCCGGTGGCCATGTGTAATCCCCCTACCGCTGCTTGAAGATACTGACTGCCCAGGA 153
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 152 ATCAAGAAGTGTGTGAAGCTCTTGGCGGATGGCTGTTCGTTTCCCCCAG 102

RESULT 31
AI874186/c
LOCUS
DEFINITION
wm50a02.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2439338 3'

similar to gb:SS8717_rnal ELAFIN PRECURSOR (HUMAN);, mRNA sequence.
AI874186
AI874186.1 GI:5548235
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 403)
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html
Insert Length: 644 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source
1..405
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2439338"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"
BASE COUNT 96 a 84 c 125 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 2,96e-29 Length: 405
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI874186 (1-405)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 375 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCTATATCTTG 316
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 315 ATCCGGTGGCCATGTGTAATCCCCCTACCGCTGCTTGAAGATACTGACTGCCCAGGA 256
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 255 ATCAAGAAGTGTGTGAAGCTCTTGGCGGATGGCTGTTCGTTTCCCCCAG 205

RESULT 32
AI137392/c
LOCUS
DEFINITION
UT-H-B11-acs-b-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
IMAGE:2715238 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 406)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA-Yes.

FEATURES

source

Location/Qualifiers

1..406
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2715238"
 /clone_lib="NCI CGAP Sub3"
 /lab_hosts="DH10B (Life Technologies)"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NCI CGAP Sub3 library is a subtracted library derived from the NCI CGAP Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLE1, NCI_CGAP_Lei2, NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_G4, NCI_CGAP_G6, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
 NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_G4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2844-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.
 TAG_LIB=NCI_CGAP_Co10
 TAG_TISSUE=Colon
 TAG_SEQ=AAACG"

BASE COUNT 95 a 126 g 99 t

ORIGIN

Alignment Scores:
 Pred. No.: 2,97e-29 Length: 406
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

09-833799-13B (1-57) x AW137392 (1-406)

QY 1 AlaGlnGluProVallyGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 366 GCGCAAGAGCAGTCAAGAGTCCAGTCTCCACTAAGCGCTGCTGCCCATTTCTTG 307
 QY 21 lleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 306 ATCCGGTGGCCCATGTTGAATCCCTAACCGCTGCTTGAAGATACCTGACTGCCAGGA 247
 QY 41 lleLysLysCysCysGlyGlySerCysGlyMetAlaCysPheValProGln 57
 Db 246 ATCAAGAAAGTGTGTGAAGAGTCTTGGCGGATGGCTGTTTCGTTCCCGAG 196
 RESULT 33
 BE715604
 LOCUS CM4-HT0744-160600-201-a08 HT0744 Homo sapiens cDNA, mRNA linear EST 12-SEP-2000
 DEFINITION BE715604
 ACCESSION BE715604
 VERSION BE715604.1 GI:10103869
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 406)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=kt2=CM4-HT0744-160600-201-a08&t3=2000-06-16&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 406.
 FEATURES
 Location/Qualifiers
 1..406
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 94 a 118 c 97 g 97 t

ORIGIN

Alignment Scores:
 Pred. No.: 2,97e-29 Length: 406
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715604 (1-406)

Qy 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 133 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCTTG 192
 |||||
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 193 ATCCGGTGCAGCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 252
 |||||
 Qy 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 253 ATCAAGAAGTCTGTGAAGGCTCTTGGGATGCGCTGTTTCGTTCCCCAG 303
 |||||

RESULT 34

AI392753/c

LOCUS AI392753 418 bp mRNA linear EST 30-MAR-1999
 DEFINITION t92303.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2109652 3'
 similar to gb:S58717_inal ELAFIN PRECURSOR (HUMAN); contains element
 HGR repetitive element ;, mRNA sequence.

ACCESSION

AI392753

VERSION

AI392753.1 GI:4222300

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 418)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,

M.D., Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 609 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 411.

Location/Qualifiers

1..418

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2109652"

/clone_lib="NCI CGAP CLL1"

/tissue_type="B-cell, chronic lymphocytic leukemia"

/lab_host="DH10B"

/note="Vector: p773D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGCGGCGCATGCTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified p773 vector.

Library is normalized, and was constructed by Bento

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

Soares and M.Fatima Bonaldo."

09-833799-13B (1-57) x AI392753 (1-418)

Qy 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
 |||||
 Db 349 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCTTG 290
 |||||
 Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 |||||
 Db 289 ATCCGGTGCAGCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 230
 |||||
 Qy 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 |||||
 Db 229 ATCAAGAAGTCTGTGAAGGCTCTTGGGATGCGCTGTTTCGTTCCCCAG 179
 |||||

RESULT 35

AA586943/c

LOCUS AA586943 419 bp mRNA linear EST 26-SEP-1997
 DEFINITION nn59a01.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089096 3'
 similar to gb:S58717_inal ELAFIN PRECURSOR (HUMAN); contains element
 HGR repetitive element ;, mRNA sequence.

ACCESSION

AA586943

VERSION

AA586943.1 GI:2397757

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 419)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 615 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 335.

Location/Qualifiers

1..419

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1089096"

/clone_lib="NCI CGAP Lar1"

/tissue_type="larynx"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: larynx; Vector: Bluescript SK-; Site 1:

EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGCGCAGG

3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

Average insert size: 0.9 kb."

92 c 135 g 89 t

103 a

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 3,09e-29

Score: 327.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

Indels: 9

Gaps: 0

09-833799-13B (1-57) x AA586943 (1-419)

Qy 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20

|||

|||

|||

|||

Db	349	GCACAAGAGCCAGTCAAAAGGTCACAGTCTCCACTAAGGCTCGCTCCGCCCATTAATCTTG	290
Qy	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	289	ATCCGGTCGCCCATGTGGAATCCCCCTAACCGCTGCTTGAAGAAGATATGACTGCCACAGA	230
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln	57
Db	229	ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTTTCGTTCCGCCAG	179
RESULT 36			
AW082097/c			
LOCUS			
DEFINITION			
ACCESSION	AW082097	422 bp	mRNA
VERSION	AW082097.1	GI:6037249	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT	98 a	88 c	129 g
ORIGIN			
Alignment Scores:			
Pred. No.:	3,11e-29	Length:	422
Score:	327.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
09-833799-13B (1-57) x AW082097 (1-422)			
Qy	1	AlaGlnGluProValLysCysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	371	GGCAAGAGCCAGTCAAAAGTCCAGTCTCCACTAAGGCTCGCTCCGCCCATTAATCTTG	312
Qy	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40

Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 273 ATCCGGTGGCGCATGTTGAATCCCGCTTAAACCGTCTGTTGAAAGATAGTACTGACTGCCAGGA 214

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 213 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGCCCTGTTTCGTTCCCGAG 163

RESULT 38
 AI283910/c 429 bp mRNA linear EST 21-DEC-1998
 LOCUS qt56a10.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:1960218 3'
 DEFINITION Similar to gb:558177_nal ELAFIN PRECURSOR (HUMAN); contains element
 THR repetitive element ; mRNA sequence.

ACCESSION AI283910
 VERSION AI283910.1 GI:3922143
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 429)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 TITLE Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
 Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert length: 626 Std Error: 0.00
 Seq primer: -40UP from Gbco
 High quality sequence stop: 428.

FEATURES
 Location/Qualifiers
 1..429
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1960218"
 /clone_lib="NCI CGAP Eso2"
 /tissue_type="squamous cell carcinoma"
 /lab_host="DH108"
 /note="Organ: esophagus; Vector: pCMV-SPORT6; Site 1: Sali
 ; Site 2: NotI; Cloned unidirectionally. Primer: Oligo
 dT. Average insert size 1.1 kb. Life Technologies catalog
 #: 11502-010"

BASE COUNT 102 a 95 c 136 g 93 t 3 others
 ORIGIN

Alignment Scores:
 Pred. No.: 3,17e-29 Length: 429
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

09-833799-13B (1-57) x AI283910 (1-429)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
 Db 350 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTGCCCGCCATTATCTTG 291

Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 290 ATCCGGTGGCGCATGTTGAATCCCGCTTAAACCGTCTGTTGAAAGATAGTACTGACTGCCAGGA 231

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 230 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGCCCTGTTTCGTTCCCGAG 180

RESULT 39
 BF836224 437 bp mRNA linear EST 13-JAN-2001
 LOCUS QV3-HT1016-171100-474-h01 HT1016 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF836224
 ACCESSION BF836224
 VERSION BF836224.1 GI:12187935
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 437)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, F.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-HT1016-171100-474-h01&t3=2000-11-17&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 13
 High quality sequence stop: 437.

FEATURES
 Location/Qualifiers
 1..437
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT1016"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 99 a 120 c 118 g 100 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,24e-29 Length: 437
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BF836224 (1-437)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
 Db 198 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTGCCCGCCATTATCTTG 257

Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 213 ATCAAGAGTCTGTGAAGGCTCTTGGGGATGCCCTGTTTCGTTCCCGAG 163

Db 258 ATCCGTCGCCATGTTGAATCCCTAACCGCTGCTTGAAGATAGTACTGACGCCAGGA 317

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||

Db 318 ATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGCGCTGTTTCGTTCCCCAG 368
|||||

RESULT 40
AW001880/c
LOCUS
DEFINITION
IMAGE:2513867.3, similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN)
) : contains element MER28 repetitive element //, mRNA sequence.
ACCESSION
AW001880
VERSION
AW001880.1 GI:5848796
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 438)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
Goldman, G.H., Carvalho, A.F., Mateukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=4t2=CM4-HT0744-160
600-201-a04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 437.
Location/Qualifiers
1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stages="Adult"
/note="Organ: head neck; Vector: puc18, Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196.716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 106 a 92 c 125 g 115 t
ORIGIN
Alignment Scores:
Pred. No.: 3,258-29 Length: 438
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x AW001880 (1-438)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 375 GCGCAAGAGCCAGTCAAGGTCAGTTTCCACTAAGCTGGCTCTCGCCCATTTATTTTG 316
|||||
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 315 ATCCGTCGCCATGTGTAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACGCCAGGA 256
|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 255 ATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGCGCTGTTTCGTTCCCCAG 205
|||||

RESULT 41
BE715601/c
LOCUS
DEFINITION
CM4-HT0744-160600-201-a04 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE715601
VERSION
BE715601.1 GI:10103866
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 438)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
Goldman, G.H., Carvalho, A.F., Mateukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=4t2=CM4-HT0744-160
600-201-a04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 437.
Location/Qualifiers
1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stages="Adult"
/note="Organ: head neck; Vector: puc18, Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196.716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 102 c 132 g 100 t
ORIGIN
Alignment Scores:
Pred. No.: 3,258-29 Length: 438
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE715601 (1-438)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 327 GCGCAAGAGCCAGTCAAGGTCAGTTTCCACTAAGCTGGCTCTCGCCCATTTATTTTG 268
|||||
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 267 ATCCGTCGCCATGTGTAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACGCCAGGA 208
|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||

Db 207 ATCAAGAGTGTCTGTAAGGCTCTTGGGGATGGCTTTTCGTTCCCCAG 157

RESULT 42
BF836743

LOCUS
CM2-HT0969-181100-509-b06 HT0969 Homo sapiens cDNA, mRNA sequence.

DEFINITION
BF836743

ACCESSION
BF836743.1 GI:12188790

VERSION
EST.

KEYWORDS
SOURCE

ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
1 (bases 1 to 441)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
20202663

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-HT0969-
181100-509-b06&t3=2000-11-18&t4=1)
Seq primer: puc18 forward
High quality sequence stop: 8
High quality sequence stop: 441.

FEATURES
source
1..441
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0969"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 94 a 131 c 109 g 107 t

ORIGIN

Alignment Scores:
Pred. No.: 3,28e-29 Length: 441
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF836743 (1-441)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
|||||

Db 148 GCGCAAGAGCCAGTCAAAAGGTCCAGTCTCCACTAAGCGCTGGCTCTGCCCCATTATCTTG 207
|||||

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

Db 208 ATCCGGTGGCCAGTGTGAATCCCGCTTGAAGATGACTGACTGCCCCAGGA 267
|||||

Qy 41 IleLysLysCysCysGlyGlySerCysGlyMetAlaCysPheValProGln 57
|||||

Db 268 ATCAAGAGTGTCTGTAAGGCTCTTGGGGATGGCTTTTCGTTCCCCAG 318
|||||

RESULT 43
AI582329

LOCUS
AI582329/c

DEFINITION
tq66g09.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2213824 3',
similar to gb:58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ;, mRNA sequence.

ACCESSION
AI582329

VERSION
AI582329.1 GI:4568226

KEYWORDS
EST.

SOURCE
human.

ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
1 (bases 1 to 448)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

TITLE
JOURNAL

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40UP from Gibco
POLYA-No.

FEATURES
Location/Qualifiers
1..448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2213824"
/clone_lib="NCI CGAP_Lul9"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 105 a 96 c 138 g 109 t

ORIGIN

Alignment Scores:
Pred. No.: 3,34e-29 Length: 448
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI582329 (1-448)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
|||||

Db 364 GCGCAAGAGCCAGTCAAAAGGTCCAGTCTCCACTAAGCGCTGGCTCTGCCCCATTATCTTG 305
|||||

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

```

Db 304 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGAGATACTGACTGCCCGGA 245
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 244 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGAG 194
|||||

RESULT 44
AI459240/C
LOCUS AI459240 453 bp mRNA linear EST 14-APR-1999
DEFINITION tk11c09.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2150704 3'
ACCESSION similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
KEYWORDS THR repetitive element ;, mRNA sequence.
SOURCE AI459240
VERSION AI459240.1 GI:4311819
EST.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 453)
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brustein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
,M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 554 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2150704"
/clone_lib="NCI-CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 107 a 96 c 137 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 3,38e-29 Length: 453
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI459240 (1-453)

QY 1 AlaGlnCluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 366 GCCAAGAGCCAGTCAAGGCTCCAGTCTCCACTAAGCCTGGCTCTGCCCATATCTTG 307
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

```

```

Db 306 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGAGATACTGACTGCCCGGA 247
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 246 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGAG 196
|||||

RESULT 45
BF824934
LOCUS BF824934 465 bp mRNA linear EST 13-JAN-2001
DEFINITION ILO-HN0039-101100-500-b06 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824934
VERSION BF824934.1 GI:12167001
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 465)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brustein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
,M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=ILO&t2=ILO-HN0039-
101100-500-b06&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 465.
FEATURES
Location/Qualifiers
1..465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head_normal; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 98 a 137 c 117 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 3,49e-29 Length: 465
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF824934 (1-465)

QY 1 AlaGlnCluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 183 GCCAAGAGCCAGTCAAGGCTCCAGTCTCCACTAAGCCTGGCTCTGCCCATATCTTG 242
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

```


Db 243 ATCCGTCGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGCTCCAGGA 302
 Qy 41 IletyLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 303 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTTTCGTTCCCCAG 353
 RESULT 46
 BF824930 466 bp mRNA linear EST 13-JAN-2001
 LOCUS ILO-HN0039-101100-500-a05 HN0039 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF824930
 ACCESSION BF824930
 VERSION BF824930.1 GI:12166995
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 466)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=ILO&t2=ILO-HN0039-101100-500-a05&t3=2000-11-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 466.

FEATURES
 source
 1..466
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 98 a 138 c 117 g 113 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.5e-29 Length: 466
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 09-833799-13B (1-57) x BF824930 (1-466)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 184 GCGCAAGACCGGTCAAGGTCCAGTCTCCACTAAGCGCTGCTCCCTCCAGCCCATATCTTTG 243

Qy 21 IletyCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 244 ATCCGTCGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGCTCCAGGA 303
 Qy 41 IletyLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 304 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTTTCGTTCCCCAG 354
 RESULT 47
 BF824935 466 bp mRNA linear EST 13-JAN-2001
 LOCUS ILO-HN0039-101100-500-c03 HN0039 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF824935
 ACCESSION BF824935
 VERSION BF824935.1 GI:12167002
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 466)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=ILO&t2=ILO-HN0039-101100-500-c03&t3=2000-11-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 466.

FEATURES
 source
 1..466
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 98 a 138 c 117 g 113 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.5e-29 Length: 466
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 09-833799-13B (1-57) x BF824935 (1-466)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 184 GCGCAAGACCGGTCAAGGTCCAGTCTCCACTAAGCGCTGCTCCCTCCAGCCCATATCTTTG 243

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 244 ATCCGGTGGCGCATGTTGAATCCCGCTAACCGCTGCTTGAAGAAGATAGTACTGCCAGGA 303

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 304 ATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCGAG 354

RESULT 48
 BF824920
 LOCUS BF824920 467 bp mRNA linear EST 13-JAN-2001
 DEFINITION ILO-HN0039-101100-500-c05 HN0039 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF824920
 VERSION BF824920.1 GI:12166980
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 467)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ILO&t2=ILO-HN0039-101100-500-c05&t3=2000-11-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 467.
 FEATURES
 source
 Location/Qualifiers
 1..467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 99 a 137 c 117 g 114 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,51e-29 Length: 467
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BF824920 (1-467)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProGlyLeu 20
 |||

Db 184 GCGCAAGAGCGGTCAAAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCCATTACTTTG 243

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 244 ATCCGGTGGCGCATGTTGAATCCCGCTAACCGCTGCTTGAAGAAGATAGTACTGCCAGGA 303

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 304 ATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCGAG 354

RESULT 49
 BF825207/c
 LOCUS BF825207 467 bp mRNA linear EST 13-JAN-2001
 DEFINITION ILO-HN0039-131100-501-d07 HN0039 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF825207
 VERSION BF825207.1 GI:12167451
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 467)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ILO&t2=ILO-HN0039-131100-501-d07&t3=2000-11-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 467.
 FEATURES
 source
 Location/Qualifiers
 1..467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 116 a 115 c 137 g 99 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,51e-29 Length: 467
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BF825207 (1-467)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProGlyLeu 20
 |||

```

|||||
Db 283 GCGAAGAGCCAGTCAAAGGTCAGTCCACTAAGCGTGGCTCTGCCCAATATCTTG 224
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 223 ATCCGGTGGCCATGTTGAATCCCTAACCGCTGCTTGAAGATACTGACTGCCCCAGGA 164
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 163 ATCAAGAAGTGTGTGAAGGCTCTTCCGGGATGCGCTGTTTCGTTCCCCAG 113

RESULT 50
BF837734/c
LOCUS BF837734 471 bp mRNA linear EST 13-JAN-2001
DEFINITION QV3-HT1016-221100-480-g03 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF837734
VERSION BF837734.1 GI:12189956
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 471)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=QV3&t2=QV3-HT1016-
221100-480-g03&t3=2000-11-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 468.
Location/Qualifiers
1. .471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 114 c 138 g 102 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3 54e-29 Length: 471
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF837734 (1-471)

|||||
Db 283 GCGAAGAGCCAGTCAAAGGTCAGTCCACTAAGCGTGGCTCTGCCCAATATCTTG 224
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 223 ATCCGGTGGCCATGTTGAATCCCTAACCGCTGCTTGAAGATACTGACTGCCCCAGGA 164
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 163 ATCAAGAAGTGTGTGAAGGCTCTTCCGGGATGCGCTGTTTCGTTCCCCAG 113

RESULT 51
BF824932
LOCUS BF824932 477 bp mRNA linear EST 13-JAN-2001
DEFINITION IL0-HN0039-101100-500-al2 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824932
VERSION BF824932.1 GI:12166998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 477)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL0&t2=IL0-HN0039-
101100-500-al2&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 476.
Location/Qualifiers
1. .477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 103 a 139 c 121 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 3 6e-29 Length: 477
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF824932 (1-477)

```

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 184 GCGCAAGAGCCAGTCAAGGTCCTCCACTAAGCGTGGCTCTCTGCCCATATATCTTG 243

Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 244 ATCCGGTGGCGCATGTTGAATCCCGCTTAACCGCTGCTTGAAGAAGATACAGTACGCCACGA 303

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 304 ATCAAGAGTCTGTGAAGGCTCTTGGCGGATGGCTGTTCTTCGTTCCCCAG 354

RESULT 52
 AA582866/c
 LOCUS
 DEFINITION
 n72a09.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089400 3'
 similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
 HGR repetitive element ;, mRNA sequence.

ACCESSION
 AA582866
 VERSION
 AA582866.1 GI:2360226
 EST.
 SOURCE
 human.

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 480)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Cloning by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
www.bio.lnlnl.gov/bbrp/image/image.html
 Insert Length: 647 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 430.

FEATURES
 Location/Qualifiers
 1..480
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1089400"
 /clone_lib="NCI CGAP Lar1"
 /tissue_type="larynx"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: larynx; Vector: Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dr. larynx. 5' adaptor sequence: 5' GAATTCGGACACGAG
 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
 Average insert size: 0.9 kb."

BASE COUNT 113 a 107 c 144 g 116 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,62e-29 Length: 480
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

09-833799-13B (1-57) x AA582866 (1-480)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
 Db 363 GCGCAAGAGCCAGTCAAGGTCCTCCACTAAGCGTGGCTCTCTGCCCATATATCTTG 304

Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 303 ATCCGGTGGCGCATGTTGAATCCCGCTTAACCGCTGCTTGAAGAAGATACAGTACGCCACGA 244

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 243 ATCAAGAGTCTGTGAAGGCTCTTGGCGGATGGCTGTTCTTCGTTCCCCAG 193

RESULT 53
 BF825204
 LOCUS
 DEFINITION
 IL0-HN0039-131100-501-a07 HN0039 Homo sapiens cDNA, mRNA sequence.

ACCESSION
 BF825204
 VERSION
 BF825204.1 GI:12167445
 EST.
 SOURCE
 human.

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 487)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

AUTHORS
 Shortgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL0&t2=IL0-HN0039-131100-501-a07&t3=2000-11-13&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 11
 High quality sequence stop: 487.

FEATURES
 Location/Qualifiers
 1..487
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 100 a 146 c 120 g 121 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,69e-29 Length: 487
 Score: 327.00 Matches: 57
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

09-833799-13B (1-57) x BF825204 (1-487)

```

Db 205 GCGCAAGAGCCAGTCAAAGGTCCTCCACTAAGCTCTGCTCTGCCCATTTATCTTG 264
|||||
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 265 ATCCGGTGGCCATGCTGAATCCCCCTAACCGCTGCTTGAAGAATATGACTGCTCCAGGA 324
|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 325 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 375
|||||

RESULT 54
BF002099/c
LOCUS
DEFINITION
799b04.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3314575 3'
similar to SW:ELAF HUMAN P19957 ELAFIN PRECURSOR ; contains element
MER28 repetitive element ;, mRNA sequence.
ACCESSION
BF002099
VERSION
BF002099.1 GI:10702374
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 490)
REFERENCE
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-femail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 421.
FEATURES
source
Location/Qualifiers
1..490
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3314575"
/tissue="NCI CGAP Col6"
/tissue_type="colon tumor, RER+"
/lab_host="DH10B"
/notes="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 119 a 113 c 146 g 112 t
ORIGIN
Alignment Scores:
Pred. No.: 3,71e-29 Length: 490
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF002099 (1-490)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||

```

```

Db 355 GCGCAAGAGCCAGTCAAAGGTCCTCCACTAAGCTCTGCTCTGCCCATTTATCTTG 296
|||||
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 295 ATCCGGTGGCCATGCTGAATCCCCCTAACCGCTGCTTGAAGAATATGACTGCTCCAGGA 236
|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 235 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 185
|||||

RESULT 55
BF837735
LOCUS
DEFINITION
QV3-HT1016-221100-480-508 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BF837735
VERSION
BF837735.1 GI:12189957
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 498)
REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.;
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
COMMENT
Contact: Simpson A.J.G
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-HT1016-221100-480-508&t3=2000-11-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 498.
FEATURES
source
Location/Qualifiers
1..498
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/notes="Organ: head neck; Vector: puc18; Site1: SmaI;
Site2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 146 c 128 g 120 t
ORIGIN
Alignment Scores:
Pred. No.: 3,78e-29 Length: 498
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF837735 (1-498)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||

```

Db	203	CGCGAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCTCGCTGCCCATTAATCTTG	262
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	263	ATCCGGTGGCCAGTGTGAATCCCTTAAACCGCTGCTTGAAGATAGTACTGCCAGGA	322
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln	57
Db	323	ATCAAGAAGTCTGTGAAGGCTCTTGCGGATGGCTGTTCGTTCCTCCAG	373
RESULT 56			
BQ941085			
LOCUS			
DEFINITION	BQ941085	522 bp mRNA linear EST 21-AUG-2002	
ACCESSION	AGENCOURT	8794612 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374832	
VERSION		5', mRNA sequence.	
KEYWORDS	BQ941085		
SOURCE	BQ941085.1	GI:22356563	
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: DCTD/BTP/Gazdar		
	cDNA Library Preparation: Rubin Laboratory		
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Distribution: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLCW2554 row: h column: 01		
	High quality sequence stop: 521.		
FEATURES	Location/Qualifiers		
source	1..522		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6374832"		
	/clone_lib="NIH_MGC_18"		
	/tissue_type="large cell carcinoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:		
	into EcoRI/XhoI sites using the following 5' adaptor:		
	into EcoRI/XhoI sites using the following 5' adaptor:		
	GGCAGCAG(G). Library constructed by Ling Hong in the		
	laboratory of Gerald M. Rubin (University of California,		
	Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and		
	Superscript II RT (Life Technologies). Note: this is a		
	NIH_MGC Library."		
BASE COUNT	124 a 155 c 121 g 122 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4e-29	Length:	522
Score:	327.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
09-833799-13B (1-57) x BQ941085 (1-522)			
Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	156	GGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATTAATCTTG	215
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40

Db	216	ATCCGGTGGCCAGTGTGAATCCCTTAAACCGCTGCTTGAAGATAGTACTGCCAGGA	275
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln	57
Db	276	ATCAAGAAGTCTGTGAAGGCTCTTGCGGATGGCTGTTCGTTCCTCCAG	326
RESULT 57			
BQ490472			
LOCUS			
DEFINITION	BQ490472	565 bp mRNA linear EST 27-MAR-2001	
ACCESSION	602519610F1	NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4637906 5',	
VERSION		mRNA sequence.	
KEYWORDS	BQ490472		
SOURCE	BQ490472.1	GI:13451982	
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: DCTD/BTP/Gazdar		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLCW1399 row: d column: 03		
	High quality sequence stop: 564.		
FEATURES	Location/Qualifiers		
source	1..565		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4637906"		
	/clone_lib="NIH_MGC_18"		
	/tissue_type="large-cell carcinoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:		
	EcoRI; cDNA made by oligo-dT priming. Directionally cloned		
	into EcoRI/XhoI sites using the following 5' adaptor:		
	GGCAGCAG(G). Library constructed by Ling Hong in the		
	laboratory of Gerald M. Rubin (University of California,		
	Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and		
	Superscript II RT (Life Technologies). Note: this is a		
	NIH_MGC Library."		
BASE COUNT	141 a 163 c 130 g 131 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4.39e-29	Length:	565
Score:	327.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0
09-833799-13B (1-57) x BQ490472 (1-565)			
Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	188	GGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATTAATCTTG	247
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	248	ATCCGGTGGCCAGTGTGAATCCCTTAAACCGCTGCTTGAAGATAGTACTGCCAGGA	307
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln	57
Db	308	ATCAAGAAGTCTGTGAAGGCTCTTGCGGATGGCTGTTCGTTCCTCCAG	358

RESULT 58
BE772161/c
LOCUS BE772161 361 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-c09 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772161
VERSION BE772161.1 GI:10225819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 361)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-c09&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 360.
FEATURES
Location/Qualifiers
1..361
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 85 c 105 g 85 t
ORIGIN

Alignment Scores:
Pred. No.: 5.89e-29 Length: 361
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE772161 (1-361)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 256 GCGAAGATCCAGTCAAGGTCCAGTCTCCACTAAGCTGGCTGCTGCCCATATCTTG 197
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 196 ATCCGGTGGCCATGTTGAATCCCCCTACCGCTGCTTGAAGATATCTGACTGCCAGGA 137
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 136 ATCAAGAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 86
RESULT 59
BE772169
LOCUS BE772169 421 bp mRNA linear EST 15-JUN-2001
DEFINITION IL3-UT0116-020201-464-F04 UT0116 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772169
VERSION BE772169.1 GI:14468596
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=IL3-UT0116-
020201-464-F04&t3=2001-02-02&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 353.
FEATURES
Location/Qualifiers
1..421
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0116"
/dev_stage="Adult"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 116 c 113 g 98 t
ORIGIN

Alignment Scores:
Pred. No.: 7.06e-29 Length: 421
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 13 Gaps: 0
09-833799-13B (1-57) x BE772169 (1-421)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 185 GCGAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCTGGCTGCTGCCCATATCTTG 244
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 245 ATCCGGTGGCCATGTTGAATCCCCCTACCGCTGCTTGAAGATATCTGACTGCCAGGA 304

```

Qy 41 IIElyLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 305 ATCAGGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCCAG 355
|||||
RESULT 60
BE711100
LOCUS RC3-HT0649-270700-012-f03 HT0649 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE711100
ACCESSION BE711100.1 GI:10099365
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 263)
REFERENCE
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zagó,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalhal,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=kt2-RC3-HT0649-270
700-012-f03&t3=2000-07-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 50
High quality sequence stop: 263.
FEATURES
Location/Qualifiers
1..263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (O.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 62 a 71 c 74 g 56 t
ORIGIN
Alignment Scores:
Pred. No.: 5,33e-29 Length: 263
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE711100 (1-263)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 57 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCAGCCCTGCTCTGCCCCATATCTTG 116
Qy 21 IIEAryCysAlaMetLeuAsnProAenArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 117 ATCCGGTGGCCCATGTGTAATCCCCCTAACCGCTGCTTTGAAAGATACTGACTGCCAGGA 176
Qy 41 IIElyLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 177 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCCAG 227
|||||
RESULT 61
AI924155 420 bp mRNA linear EST 08-MAR-2000
wn64c08.x1 NCI CGAP Lul19 Homo sapiens cDNA clone IMAGE:2450222 3'
similar to gb:S58717.rnal ELAFIN PRECURSOR (HUMAN); contains element
PTR5 repetitive element ;, mRNA sequence.
AI924155
ACCESSION AI924155.1 GI:5660119
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 420)
REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Cloning Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 511 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2450222"
/clone_lib="NCI CGAP_Lul19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from a
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 102 a 89 c 130 g 99 t
ORIGIN
Alignment Scores:
Pred. No.: 9,26e-29 Length: 420
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AI924155 (1-420)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 358 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCAGCCCTGCTCTGCCCCATATCTTG 299
|||||

```



```

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 298 ATCCGGTGCACCATGTTGAATCCCTTAAACCGCTGCTTGAAGATACCTACTGCCAGGA 239
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 238 ATCAAGAAGTCTGTAAAGGCTCTTGGGGATGCCCTGTTTCGTTCCCCAG 188

RESULT 62
BF835679
LOCUS QV3-HT1016-151100-463-c10 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF835679
VERSION BF835679.1 GI:12186947
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 447)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-HT1016-
151100-463-c10&t3=2000-11-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 447.
FEATURES
source
1. .447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 96 a 128 c 117 g 106 t
ORIGIN
Alignment Scores:
Pred. No.: 9-97e-29 Length: 447
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF835679 (1-447)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
||||:|||||

```

```

Db 189 GCGCAGAGCCAGTCAAAGGCTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 248
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 249 ATCCGGTGCACCATGTTGAATCCCTTAAACCGCTGCTTGAAGATACCTACTGCCAGGA 308
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 309 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGCCCTGTTTCGTTCCCCAG 359

RESULT 63
AW845140/c
LOCUS AW845140 453 bp mRNA linear EST 19-MAY-2000
DEFINITION ILO-CT0008-140599-007 CT0008 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW845140
VERSION AW845140.1 GI:7940657
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 453)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=ILO-CT0008-140
599-007&t3=1999-05-14&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 454.
FEATURES
source
1. .453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0008"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: Smal; Site 2:
Smal; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 106 a 110 c 139 g 98 t
ORIGIN
Alignment Scores:
Pred. No.: 1.01e-28 Length: 453
Score: 323.00 Matches: 56
Percent Similarity: 98.25% Conservatives: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AW845140 (1-453)

```

```

QY 1 AlaGluProVallysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 307 GCGAAGAGCCAGTCAAGAGTCCAGCTCCACTAGAGCTGGCTCTGCCCATTAATCTTG 248
QY 21 IleArGysAlaMetLeuAsnProProAsnArGysLeuLysAspThrAspCysProGly 40
Db 247 ATCCGGTGGCCATGTTGAATCCCTCAACCGCTGCTTGAAGAGATACACTGCGCCAGGA 188
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 187 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTGTTCCCGG 137

RESULT 64
LOCUS AW845142/c 455 bp mRNA linear EST 19-MAY-2000
DEFINITION ILO-CT0008-140599-001 CT0008 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW845142
VERSION AW845142.1 GI:7940659
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 455)
AUTHORS Dias Neto,E., Garcia Correa,R., Vertovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 2020263
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=ILO-CT0008-140
599-001&t3=1999-05-14&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 455.
Location/Qualifiers
1..455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0008"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site: 1: Smal; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 108 a 116 c 134 g 97 t
ORIGIN
Alignment Scores:
Pred. No.: 1,02e-28 Length: 455
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x AW845142 (1-455)

```

```

QY 1 AlaGluProVallysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 282 GCGAAGAGCCAGTCAAGAGTCCAGTCTCCACTAGAGCTGGCTCTGCCCATTAATCTTG 223
QY 21 IleArGysAlaMetLeuAsnProProAsnArGysLeuLysAspThrAspCysProGly 40
Db 222 ATCCGGTGGCCATGTTGAATCCCTCAACCGCTGCTTGAAGAGATACACTGACTGCCAGGA 163
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 162 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTGTTCCCGG 112

RESULT 65
LOCUS AI858070/c 444 bp mRNA linear EST 21-DEC-1999
DEFINITION wj70a03.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2408140 3'
similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ;, mRNA sequence.
ACCESSION AI858070
VERSION AI858070.1 GI:5511686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA library preparation: M. Bento Soares, Ph.D.
DNA sequencing by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 551 Std Error: 0.00
Seq primer: -400P from Gibco.
Location/Qualifiers
1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2408140"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pVT3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pVT3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 108 a 96 c 134 g 105 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,3e-28 Length: 444
Score: 322.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 98.47% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AI858070 (1-444)

```

```

09-833799-13B (1-57) x A1858070 (1-444)
Qy 1 AlaGlnGluProVallySerThryLysProGlySerCysProIleIleLeu 20
Db 360 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCANTAAAGCTGGCTCCTCCCATTAATCTTG 301
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 300 ATCCGGTGGCGCATGTTGAATCCCCCTTAACCGCTGCTTGAAGATACTAGACTGCCAGGA 241
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 240 ATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCCCTGTTTCGTTCCCCAG 190

RESULT 66
BE715619/c 342 bp mRNA linear EST 12-SEP-2000
LOCUS CM4-HT0744-160600-201-f01 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715619
VERSION BE715619.1 GI:10103884
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 342)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=kt2=CM4-HT0744-160
600-201-f01&kt3=2000-06-16&kt4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 342.
Location/Qualifiers
1..342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 78 a 87 c 93 g 83 t 1 others
ORIGIN
Alignment Scores: 1-26e-28 Length: 342
Pred. No.: 321.00 Matches: 56
Score: 98.25%
Percent Similarity: 98.25%
Best Local Similarity: 98.25%
Conservative: 0
Mismatch: 1

09-833799-13B (1-57) x BE715619 (1-342)
Qy 1 AlaGlnGluProVallySerThryLysProGlySerCysProIleIleLeu 20
Db 234 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCAAAGCTGGCTCCTCCCATTAATCTTG 175
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 174 ATCCGGTGGCGCACGTTGAATCCCCCTTAACCGCTGCTTGAAGATACTAGACTGCCAGGA 115
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 114 ATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCCCTGTTTCGTTCCCCAG 64

RESULT 67
BE715619/c 374 bp mRNA linear EST 13-JAN-2001
LOCUS QV3-HT1016-171100-474-c05 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715619
VERSION BE715619.1 GI:12187891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 374)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&kt2=QV3-HT1016-
171100-474-c05&kt3=2000-11-17&kt4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 374.
Location/Qualifiers
1..374
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 84 a 98 c 102 g 90 t
ORIGIN
Alignment Scores: 1.4e-28 Length: 374
Pred. No.: 321.00 Matches: 56
Score: 98.25%
Percent Similarity: 98.25%
Best Local Similarity: 98.25%
Conservative: 0
Mismatch: 1

```

```

Percent Similarity: 98.25%      Conservative: 0
Best Local Similarity: 98.25%    Mismatches: 1
Query Match: 98.17%             Indels: 0
DB: 12                          Gaps: 0

09-833799-13B (1-57) x BF836201 (1-374)

QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 169 GGCAGAGCCAGTCAAGGTCAGTCCACTAAGCTGGCTCTGCCCATTTATCTTG 228

QY 21 IleArgCysAlaMetLeuAsnProProhenArgCysLeuLysAspThrAspCysProGly 40
Db 229 ATCCGGTGGCCATGTTGAATCCCCCTTAAGCGCTGCTTGAAGAATACTGACTGCCAGGA 288

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 289 ATCAAGAAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCCAG 339

RESULT 68
BF825196 501 bp mRNA linear EST 13-JAN-2001
LOCUS ILO-HN0039-131100-501-d10 HN0039 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF825196
ACCESSION BF825196.1 GI:12167431
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=ILO&t2=ILO-HN0039-
131100-501-d10&t3=2000-11-13&t4=1)
Seq primer: puc 18 forward
High quality sequence scop: 501.
Location/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 109 a 146 c 127 g 119 t
ORIGIN

FEATURES
source
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 109 a 146 c 127 g 119 t
ORIGIN

Score: 321.00      Matches: 56
Percent Similarity: 98.25%      Conservative: 0
Best Local Similarity: 98.25%    Mismatches: 1
Query Match: 98.17%             Indels: 0
DB: 12                          Gaps: 0

09-833799-13B (1-57) x BF825196 (1-501)

QY 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 202 GGCAGAGCCAGTCAAGGTCAGTCCACTAAGCTGGCTCTGCCCATTTATCTTG 261

QY 21 IleArgCysAlaMetLeuAsnProProhenArgCysLeuLysAspThrAspCysProGly 40
Db 262 ATCCGGTGGCCATGTTGAATCCCCCTTAACCGCTGCTTGAAGAATACTGACTGCCAGGA 321

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 322 ATCAAGAAAGTGTGTGAAGGCTCTTGGCGGATGGACTGTTTCGTACCCAG 372

RESULT 69
BF711083/c 316 bp mRNA linear EST 12-SEP-2000
LOCUS RC3-HT0649-270700-012-f06 HT0649 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF711083
ACCESSION BF711083
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 316)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC3-HT0649-270
700-012-f06&t3=2000-07-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 316.
Location/Qualifiers
1..316
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 68 a 73 c 114 g 61 t
ORIGIN

TITLE
JOURNAL
MEDLINE
COMMENT

```

```

Alignment Scores:
Pred. No.: 1.51e-28 Length: 316
Score: 320.00 Matches: 55
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.49% Mismatches: 0
Query Match: 97.86% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE711083 (1-316)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
Db 289 GCGAAGAGCCACCAAGGTCAGTCTCCACTAAGCTGGCTCTCCCTCCCATATCTTG 230
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 229 ATCCGGTGGCGCATGTTGAATCCCTTAACCGCTGCTTGAAGATGACTGATGCTGCCAGGA 170
QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 169 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGT 119

RESULT 70
AW844969
LOCUS AW844969 344 bp mRNA linear EST 19-MAY-2000
DEFINITION MR0-CT0006-280599-008 CT0006 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW844969
VERSION AW844969.1 GI:7940577
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 344)
Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=MR0-CT0006-280
599-008&t3=1999-05-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 345.
Location/Qualifiers
1. .344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0006"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification. Reverse transcription of low
stringency conditions."
BASE COUNT 82 a 89 c 88 g 85 t

ORIGIN
Alignment Scores:
Pred. No.: 1.66e-28 Length: 344
Score: 320.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 97.86% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AW844969 (1-344)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
Db 174 GCGAGGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTCCCTCCCATATCTTG 233
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 234 ATCCGGTGGCGCATGTTGAATCCCTTAACCGCTGCTTGAAGATGACTGATGCTGCCAGGA 293
QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 294 ATCAAGAAGTCTGTGAAGGCTCTTGGCATATGGCTGTTTCGTTCCCGCAG 344

RESULT 71
BM768573
LOCUS BM768573 428 bp mRNA linear EST 04-MAR-2002
DEFINITION K-EST0051573 S14K402 Homo sapiens cDNA clone S14K402-6-G01 5', mRNA
sequence.
ACCESSION BM768573
VERSION BM768573.1 GI:19098188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 428)
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 6 row: G column: 01
High quality sequence stop: 428.
Location/Qualifiers
1. .428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="S14K402-6-G01"
/clone_lib="S14K402"
/cell_line="K402"
/lab_host="Top10P"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: ECORI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P' by electroporation method.
FEATURES
source

```

The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 101 a 120 c 108 g 99 t
ORIGIN

Alignment Scores:

Pred. No.: 2.15e-28 Length: 428
Score: 320.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 97.86% Indels: 0
DB: 14 Gaps: 0

09-833799-13B (1-57) x BM768573 (1-428)

QY 1 AlaGlnGluProVallySGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 200 GCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCATATATCTTG 259
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 260 ATCCCGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 319
QY 41 IleLysLysCysGlyGlySerCysGlyMetAlaCysPheValProGln 57
Db 320 ATCAAGAAATGCTGTGAGGCTCTTGGCGGATGGCTGTTTCGTTCCCCAG 370

RESULT 72

Bi021805/c

LOCUS Bi021805 394 bp mRNA linear EST 14-JUN-2001

DEFINITION CM3-MT0316-170101-673-e02 MT0316 Homo sapiens cDNA, mRNA sequence.

ACCESSION Bi021805

VERSION Bi021805.1 GI:14428435

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 394)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

2020263

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM3&t2=CM3-MT0316-

170101-673-e02&t3=2001-01-17&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 394.

Location/Qualifiers

1..394

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="MT0316"

/dev_stage="Adult"

/note="Organ: marrow; Vector: puc18; Site:1: SmaI; Site:2:

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196

,716 - Ludwig Institute for Cancer Research) profiles

into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 91 a 106 c 108 g 89 t
ORIGIN

Alignment Scores:

Pred. No.: 2.57e-28 Length: 394
Score: 319.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 97.55% Indels: 0
DB: 13 Gaps: 0

09-833799-13B (1-57) x Bi021805 (1-394)

QY 1 AlaGlnGluProVallySGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 229 GCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCCCATATATCTTG 170

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 169 ATCCGTTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 110

QY 41 IleLysLysCysGlyGlySerCysGlyMetAlaCysPheValProGln 57
Db 109 ATCAAGAGTGTGTGAAGCTCTTGGCGGATGGCTGTTTCGTTCCCCAG 59

RESULT 73

AA586983/c

LOCUS AA586983

DEFINITION nm9e10.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089162 3'

similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element

THR repetitive element ;, mRNA sequence.

ACCESSION AA586983

VERSION AA586983.1 GI:2397797

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 386)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

DNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 642 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 241.

Location/Qualifiers

1..386

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1089162"

/clone_lib="NCI CGAP Lar1"

/tissue_type="larynx"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: larynx; Vector: Bluescript SK-; Site:1:

EOCR; Site:2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGAG

3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

Average insert size: 0.9 kb."

BASE COUNT 86 a 75 c 121 g 104 t

FEATURES
source

```

ORIGIN
Alignment Scores:
Pred. No.: 3.3e-28 Length: 386
Score: 318.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.25% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AA586983 (1-386)

Qy 3 GluProValLysGlyProValSerThrLysProGlySerCysProLleLeuArg 22
Db 384 GAGCAGTCAAGTCCAGTCTCCACTAAGCCTGGCTCTCCGCCCATTAATCTTGATCCGG 325

Qy 23 CysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGlyLleLys 42
Db 324 TGCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCAGGAATCAAG 265

Qy 43 LysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 264 AAGTCTGTGAAGCTCTTGGCGGATGCGCTGTTTCGTTCCCCAG 220

RESULT 74
BF736734/c
LOCUS BF736734 419 bp mRNA linear EST 10-JAN-2001
DEFINITION PM1-KT0004-241100-001-c04 KT0004 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF736734
VERSION BF736734.1 GI:12063408
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 419)
Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,D.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM1&t2=PM1-KT0004-
241100-001-c04&t3=2000-11-24&t4=1)
Query primer: puc 18 forward
High quality sequence stop: 419.
Location/Qualifiers
1..419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="KT0004"
/dev_stage="Adult"
/note="Organ: bladder tumor; Vector: puc18; Site 1: Smal;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 136,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

FEATURES
source
1..423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Lul9"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p773 vector. Library went through one round of

BASE COUNT 99 a 103 c 126 g 91 t
ORIGIN
Alignment Scores:
Pred. No.: 3.63e-28 Length: 419
Score: 318.00 Matches: 55
Percent Similarity: 98.25% Conservative: 1
Best Local Similarity: 96.49% Mismatches: 1
Query Match: 97.25% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF736734 (1-419)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeu 20
Db 285 GGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAACCTTG 226

Qy 21 IleArgCysAlaMetLeuAenProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 225 ATCCGGTGGCGGCTGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCAGGA 166

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 165 ATCAAGAAAGTCTGTGAAGGCTCTTGGCGGATGCGCTGTTTCGTTCCCCAG 115

RESULT 75
AI830872/c
LOCUS AI830872 423 bp mRNA linear EST 21-DEC-1999
DEFINITION wj6ic08.xl NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2407310 3'
similar to gb:S58717.rnal ELAPIN PRECURSOR (HUMAN); contains element
MER22 repetitive element ;, mRNA sequence.
ACCESSION AI830872
VERSION AI830872.1 GI:5451543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
NCT-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 534 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Lul9"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p773 vector. Library went through one round of

```

normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 106 a 91 c 125 g 101 t
ORIGIN

Alignment Scores:

Pred. No.: 3.68e-28 Length: 423
Score: 318.00 Matches: 55
Percent Similarity: 98.25% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 1
Query Match: 97.25% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x A1830872 (1-423)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 359 GCGCAAGAGCCAGTCAAGGTCCAGTCCCAATAAGCCTGGCTCTGCCCATTAICTTG 300
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 299 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGGTGTTTGAAGAAGAACTGACTGCCCCAGGA 240
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 239 ATCAGAGAGTGCTGTGAGGTTCTTTCGGGGATGGCCTGTTTCGTTCCCCAG 189

Search completed: February 15, 2003, 21:12:55
Job time : 1467 secs

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 20:49:11 ; Search time 108 Seconds
(without alignments)
268.810 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvtbpgscpiil.....cpgikkcccgscgmactvfpq 57

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 424239 seqs, 254661826 residues
Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09833799/runat_12022003_120044_23813/app_query.fasta_1.199
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09833799 @CGN_1_1 80 @runat_12022003_120044_23813
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	327	100.0	480	10	US-09-920-300A-1393
C 2	327	100.0	480	12	US-10-033-528-1393
C 3	327	100.0	2309	10	US-09-954-456-1126
4	327	100.0	2309	10	US-09-954-456-1798

5	143	43.7	594	10	US-09-964-824A-582	Sequence 582, App
6	143	43.7	594	10	US-09-954-456-1989	Sequence 1989, App
7	143	43.7	594	10	US-09-865-812-1	Sequence 1, Appl
8	112	34.3	762	9	US-09-992-598-344	Sequence 344, App
9	112	34.3	762	9	US-09-989-293A-344	Sequence 344, App
10	112	34.3	762	9	US-09-989-735-344	Sequence 344, App
11	112	34.3	762	9	US-09-990-444-344	Sequence 344, App
12	112	34.3	762	9	US-09-989-730-344	Sequence 344, App
13	112	34.3	762	9	US-09-990-436-344	Sequence 344, App
14	112	34.3	762	9	US-09-991-181-344	Sequence 344, App
15	112	34.3	762	9	US-09-993-687-344	Sequence 344, App
16	112	34.3	762	9	US-09-989-734-344	Sequence 344, App
17	112	34.3	762	9	US-09-997-653-344	Sequence 344, App
18	112	34.3	762	9	US-09-993-667-344	Sequence 344, App
19	112	34.3	762	9	US-09-990-438-344	Sequence 344, App
20	112	34.3	762	9	US-09-990-562-344	Sequence 344, App
21	112	34.3	762	9	US-09-997-428-344	Sequence 344, App
22	112	34.3	762	9	US-09-997-666-344	Sequence 344, App
23	112	34.3	762	10	US-09-989-722-344	Sequence 344, App
24	112	34.3	762	10	US-09-989-723-344	Sequence 344, App
25	112	34.3	762	10	US-09-989-279-344	Sequence 344, App
26	112	34.3	762	10	US-09-989-727-344	Sequence 344, App
27	112	34.3	762	10	US-09-989-731-344	Sequence 344, App
28	112	34.3	762	10	US-09-989-732-344	Sequence 344, App
29	112	34.3	762	10	US-09-991-073-344	Sequence 344, App
30	112	34.3	762	10	US-09-990-442-344	Sequence 344, App
31	112	34.3	762	10	US-09-991-163-344	Sequence 344, App
32	112	34.3	762	10	US-09-993-604-344	Sequence 344, App
33	112	34.3	762	10	US-09-990-456-344	Sequence 344, App
34	112	34.3	762	10	US-09-989-721-344	Sequence 495, App
35	107	32.7	411	10	US-09-964-824A-495	Sequence 1987, App
36	107	32.7	411	10	US-09-954-456-1987	Sequence 3, Appl
37	107	32.7	411	10	US-09-967-768A-3	Sequence 377, App
38	105	32.1	292	10	US-09-884-441-377	Sequence 589, App
39	98	30.0	724	9	US-10-152-661-589	Sequence 214, App
40	97	29.7	583	10	US-09-967-768A-214	Sequence 1629, App
41	95	29.1	469	10	US-09-864-761-1629	Sequence 57, Appl
42	94.5	28.9	228	10	US-09-790-264-57	Sequence 29, Appl
43	94.5	28.9	437	10	US-09-823-038A-29	Sequence 55, Appl
44	94.5	28.9	471	10	US-09-790-264-55	Sequence 421, App
45	93	28.4	622	10	US-09-925-300-421	Sequence 41, Appl
46	93	28.4	643	9	US-09-924-340-41	Sequence 41, Appl
47	93	28.4	643	9	US-09-992-600A-41	Sequence 57, Appl
48	93	28.4	783	9	US-10-093-766-57	Sequence 319, App
49	92.5	28.3	6314	10	US-09-954-456-319	Sequence 1230, App
50	92.5	28.3	6314	10	US-09-954-456-1230	Sequence 34, Appl
51	85	26.0	753	9	US-09-852-797-34	Sequence 34, Appl
52	85	26.0	753	10	US-09-853-161-34	Sequence 50, Appl
53	85	26.0	753	10	US-09-852-659A-34	Sequence 50, Appl
54	85	26.0	783	9	US-09-852-797-50	Sequence 121, App
55	85	26.0	783	10	US-09-853-161-50	Sequence 747, App
56	85	26.0	783	10	US-09-852-659A-50	Sequence 1, Appl
57	80.5	24.6	354	10	US-09-852-659A-121	Sequence 240, App
58	78.5	24.0	334	10	US-09-833-381-747	Sequence 249, App
59	74.5	22.8	2082	10	US-09-819-136-1	Sequence 750, App
60	73.5	22.5	243	10	US-09-894-882-240	Sequence 2681, App
61	73.5	22.5	243	10	US-09-894-882-249	Sequence 1, Appl
62	71	21.7	3836	12	US-10-044-090-750	Sequence 1387, A
63	70.5	21.6	2937	10	US-10-044-090-750	Sequence 198, App
64	69.5	21.3	753	10	US-09-974-300-2288	Sequence 12659, A
65	69.5	21.3	465237	10	US-09-910-943-661	Sequence 225, App
66	69	21.1	148	10	US-09-933-267A-1	Sequence 6, Appl
67	69	21.1	403	10	US-09-864-761-18387	Sequence 702, App
68	68	20.8	511	10	US-09-884-441-198	Sequence 46, Appl
69	67.5	20.6	2108	10	US-09-864-761-12659	Sequence 47, Appl
70	67	20.5	302250	10	US-09-962-832-225	Sequence 243, App
71	66	20.2	2853	10	US-09-962-832-154	Sequence 255, App
72	66	20.2	4656	10	US-09-815-242-7863	
73	66	20.2	10172	9	US-09-891-216-6	
74	66	20.2	10284	9	US-09-954-456-702	
75	66	20.2	10284	12	US-09-974-298-46	
76	64.5	19.7	243	10	US-10-044-090-47	
77	64.5	19.7	243	10	US-09-894-882-243	
					US-09-894-882-255	

US-09-954-456-1126

; Sequence 1126, Application US/09954456

Patent No. US20020115057A1
: GENERAL INFORMATION:

APPLICANT: Young

TITLE OF INVENTION: pro

TITLE OF INVENTION: _____

FILE REFERENCE: 689290-

[illegible]

CURRENT FILING DATE: 2/2/2011
PRIOR APPLICATION NUMBER: 2009/0100000

PRIOR FILING DATE: 2000

PRIOR APPLICATION NUMBER

PRIOR FILING DATE: 2000

PRIOR FILING DATE: 2000

PRIOR APPLICATION NUMBER: 2000

PRIOR FILING DATE: 2000

PRIOR APPLICATION NUMBER

PRIOR FILING DATE: 2000-
PRIOR APPLICATION NUMBER:

PRIOR FILING DATE: 2000-

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

100


```
Db 298 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAGCGT 357
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 405
RESULT 6
US-09-954-456-1989
; Sequence 1989, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR FILING DATE: US/60/233,617
; PRIOR APPLICATION NUMBER: 2000-09-18
; PRIOR FILING DATE: US/60/234,052
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1989
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1989
Alignment Scores:
Pred. No.: 1.7e-09
Score: 143.00
Percent Similarity: 55.36%
Best Local Similarity: 48.21%
Query Match: 43.73%
DB: 10
Mismatches: 23
Indels: 2
Gaps: 1
09-833799-13B (1-57) x US-09-954-456-1989 (1-594)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 238 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGCAGTACTTAT 297
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 298 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAGCGT 357
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 405
RESULT 7
US-09-865-812-1
; Sequence 1, Application US/09865812
; Patent No. US20020115626A1
; GENERAL INFORMATION:
; APPLICANT: Rastelli, Luca
; APPLICANT: Smithson, Glenna
; TITLE OF INVENTION: Method of Detecting Inflammatory Lung Disorders
; FILE REFERENCE: 21402-018 US
; CURRENT APPLICATION NUMBER: US/09/865,812
; CURRENT FILING DATE: 2001-05-28
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-865-812-1
Alignment Scores:
Pred. No.: 1.7e-09
Score: 143.00
Percent Similarity: 55.36%
Best Local Similarity: 48.21%
Query Match: 43.73%
DB: 10
Mismatches: 23
Indels: 2
Gaps: 1
09-833799-13B (1-57) x US-09-865-812-1 (1-594)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 238 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGCAGTACTTAT 297
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 298 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAGCGT 357
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 405
RESULT 8
US-09-992-598-344
; Sequence 344, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gottsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
```

[illegible]

PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.:
Score: 1.9e-05
Percent Similarity: 112.00
Best Local Similarity: 53.57%
Query Match: 42.86%
DB: 34.25%
Indels: 9
Gaps: 2

09-833799-13B (1-57) x US-09-992-598-344 (1-762)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCACGCTGACAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGATCTCTCC---CAGTGTACACAGACAGGACTGTCTGGG 175
Qy 41 IleLysLysCysGlyGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAGAGTGTGTTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCT 223

RESULT 9
US-09-989-293A-344
Sequence 344, Application US/09989293A
Patent No. US20020177164A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC66
CURRENT APPLICATION NUMBER: US/09/989,293A
PRIOR FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861

PRIOR APPLICATION NUMBER: 60/090695	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090696	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090862	
PRIOR FILING DATE: 1998-06-26	
PRIOR APPLICATION NUMBER: 60/090863	
PRIOR FILING DATE: 1998-06-26	
PRIOR APPLICATION NUMBER: 60/091360	
PRIOR FILING DATE: 1998-07-01	
PRIOR APPLICATION NUMBER: 60/091478	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091544	
PRIOR FILING DATE: 1998-07-01	
PRIOR APPLICATION NUMBER: 60/091519	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091626	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091633	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091978	
PRIOR FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: 60/091982	
PRIOR FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: 60/092182	
PRIOR FILING DATE: 1998-07-09	

Alignment Scores:	1.9e-05	Length:	762
Pred. No.:	112.00	Matches:	24
Score:	53.57%	Conservative:	6
Percent Similarity:	42.86%	Mismatches:	24
Best Local Similarity:	34.25%	Indels:	2
Query Match:	9	Gaps:	2

DB: 200 202A-244 (1-762)

09-833799-13B (1-57) x US-09-369-223 (1-57)

QY	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
		:::	
Db	62	GCTGTGGAAAGGATTAAAGAGGTATAGAG--AAAGCAGGGGTTGCCACGCTGACAAC	118
QY	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40
		:::	
Db	119	GTACGCTGTTTCAAGTCCGATCTCTCC--CAGTGTCAACAGACAGGATGTCCTGGG	175
QY	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
		:::	
Db	176	GAAGAAGAGTGTTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT	223

RESULT 10
US-09-989-735-344 Application US/09989735

Sequence 344, App. 1
Publication No. US20020193239A1

GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnuyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/088876	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/089105	
PRIOR FILING DATE: 1998-06-12	
PRIOR APPLICATION NUMBER: 60/089440	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089512	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089514	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089532	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089538	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089598	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089599	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089600	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089653	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089801	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089907	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089908	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089947	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089948	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089952	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/090246	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090252	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090254	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090349	
PRIOR FILING DATE: 1998-06-23	
PRIOR APPLICATION NUMBER: 60/090355	
PRIOR FILING DATE: 1998-06-23	
PRIOR APPLICATION NUMBER: 60/090429	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090431	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090435	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090444	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090445	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090472	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090535	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090540	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090542	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090557	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090676	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090678	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090690	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090694	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090694	
PRIOR FILING DATE: 1998-06-25	

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC61
CURRENT FILING DATE: 2001-11-19
PRIORITY APPLICATION NUMBER: US/09/989,735
PRIORITY FILING DATE: 2001-11-19
PRIORITY APPLICATION NUMBER: 60/049787
PRIORITY FILING DATE: 1997-06-16
PRIORITY APPLICATION NUMBER: 60/062250
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/065186
PRIORITY FILING DATE: 1997-11-12
PRIORITY APPLICATION NUMBER: 60/065311
PRIORITY FILING DATE: 1997-11-13
PRIORITY APPLICATION NUMBER: 60/066770
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/075945
PRIORITY FILING DATE: 1998-02-25
PRIORITY APPLICATION NUMBER: 60/078910
PRIORITY FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: 60/083322
PRIORITY FILING DATE: 1998-04-28
PRIORITY APPLICATION NUMBER: 60/084600
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/087106
PRIORITY FILING DATE: 1998-05-28
PRIORITY APPLICATION NUMBER: 60/087607
PRIORITY FILING DATE: 1998-06-02
PRIORITY APPLICATION NUMBER: 60/087609
PRIORITY FILING DATE: 1998-06-02
PRIORITY APPLICATION NUMBER: 60/087759
PRIORITY FILING DATE: 1998-06-02
PRIORITY APPLICATION NUMBER: 60/087827
PRIORITY FILING DATE: 1998-06-03
PRIORITY APPLICATION NUMBER: 60/088021
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088025
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088026
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088028
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088029
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088030
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088033
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088326
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088167
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088202
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088212
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088217
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088655
PRIORITY FILING DATE: 1998-06-09
PRIORITY APPLICATION NUMBER: 60/088734
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088738
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088742
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088810
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088824

PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088826
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088858
PRIORITY FILING DATE: 1998-06-11
PRIORITY APPLICATION NUMBER: 60/088861
PRIORITY FILING DATE: 1998-06-11
PRIORITY APPLICATION NUMBER: 60/088876
PRIORITY FILING DATE: 1998-06-11
PRIORITY APPLICATION NUMBER: 60/089105
PRIORITY FILING DATE: 1998-06-12
PRIORITY APPLICATION NUMBER: 60/089440
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: 60/089512
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: 60/089514
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: 60/089532
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089538
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089598
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089599
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089600
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089653
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089801
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/089907
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/089908
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/089947
PRIORITY FILING DATE: 1998-06-19
PRIORITY APPLICATION NUMBER: 60/089948
PRIORITY FILING DATE: 1998-06-19
PRIORITY APPLICATION NUMBER: 60/089952
PRIORITY FILING DATE: 1998-06-19
PRIORITY APPLICATION NUMBER: 60/090246
PRIORITY FILING DATE: 1998-06-22
PRIORITY APPLICATION NUMBER: 60/090252
PRIORITY FILING DATE: 1998-06-22
PRIORITY APPLICATION NUMBER: 60/090254
PRIORITY FILING DATE: 1998-06-22
PRIORITY APPLICATION NUMBER: 60/090349
PRIORITY FILING DATE: 1998-06-23
PRIORITY APPLICATION NUMBER: 60/090355
PRIORITY FILING DATE: 1998-06-23
PRIORITY APPLICATION NUMBER: 60/090429
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090431
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090435
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090444
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090445
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090472
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090535
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090540
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090542
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090557
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090676
PRIORITY FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1,98-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Watch: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-989-735-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGATCCCTCCC---CAGTGTCAACAGACAGAGGACTGTCTGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCT 223

RESULT 11

US-09-990-444-344

; Sequence 344, Application US/09990444

; Publication No. US20020193300A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: Acids Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/990,444
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738

1	PRIOR FILING DATE: 1998-06-10
2	PRIOR APPLICATION NUMBER: 60/088742
3	PRIOR FILING DATE: 1998-06-10
4	PRIOR APPLICATION NUMBER: 60/088810
5	PRIOR FILING DATE: 1998-06-10
6	PRIOR APPLICATION NUMBER: 60/088810
7	PRIOR FILING DATE: 1998-06-10
8	PRIOR APPLICATION NUMBER: 60/088824
9	PRIOR FILING DATE: 1998-06-10
10	PRIOR APPLICATION NUMBER: 60/088826
11	PRIOR FILING DATE: 1998-06-10
12	PRIOR APPLICATION NUMBER: 60/088858
13	PRIOR FILING DATE: 1998-06-11
14	PRIOR APPLICATION NUMBER: 60/088861
15	PRIOR FILING DATE: 1998-06-11
16	PRIOR APPLICATION NUMBER: 60/088876
17	PRIOR FILING DATE: 1998-06-11
18	PRIOR APPLICATION NUMBER: 60/089105
19	PRIOR FILING DATE: 1998-06-12
20	PRIOR APPLICATION NUMBER: 60/089440
21	PRIOR FILING DATE: 1998-06-16
22	PRIOR APPLICATION NUMBER: 60/089512
23	PRIOR FILING DATE: 1998-06-16
24	PRIOR APPLICATION NUMBER: 60/089514
25	PRIOR FILING DATE: 1998-06-16
26	PRIOR APPLICATION NUMBER: 60/089532
27	PRIOR FILING DATE: 1998-06-17
28	PRIOR APPLICATION NUMBER: 60/089538
29	PRIOR FILING DATE: 1998-06-17
30	PRIOR APPLICATION NUMBER: 60/089598
31	PRIOR FILING DATE: 1998-06-17
32	PRIOR APPLICATION NUMBER: 60/089599
33	PRIOR FILING DATE: 1998-06-17
34	PRIOR APPLICATION NUMBER: 60/089600
35	PRIOR FILING DATE: 1998-06-17
36	PRIOR APPLICATION NUMBER: 60/089653
37	PRIOR FILING DATE: 1998-06-17
38	PRIOR APPLICATION NUMBER: 60/089801
39	PRIOR FILING DATE: 1998-06-18
40	PRIOR APPLICATION NUMBER: 60/089907
41	PRIOR FILING DATE: 1998-06-18
42	PRIOR APPLICATION NUMBER: 60/089908
43	PRIOR FILING DATE: 1998-06-18
44	PRIOR APPLICATION NUMBER: 60/089947
45	PRIOR FILING DATE: 1998-06-19
46	PRIOR APPLICATION NUMBER: 60/089948
47	PRIOR FILING DATE: 1998-06-19
48	PRIOR APPLICATION NUMBER: 60/089952
49	PRIOR FILING DATE: 1998-06-19
50	PRIOR APPLICATION NUMBER: 60/090246
51	PRIOR FILING DATE: 1998-06-22
52	PRIOR APPLICATION NUMBER: 60/090252
53	PRIOR FILING DATE: 1998-06-22
54	PRIOR APPLICATION NUMBER: 60/090254
55	PRIOR FILING DATE: 1998-06-22
56	PRIOR APPLICATION NUMBER: 60/090349
57	PRIOR FILING DATE: 1998-06-23
58	PRIOR APPLICATION NUMBER: 60/090355
59	PRIOR FILING DATE: 1998-06-23
60	PRIOR APPLICATION NUMBER: 60/090429
61	PRIOR FILING DATE: 1998-06-24
62	PRIOR APPLICATION NUMBER: 60/090431
63	PRIOR FILING DATE: 1998-06-24
64	PRIOR APPLICATION NUMBER: 60/090435
65	PRIOR FILING DATE: 1998-06-24
66	PRIOR APPLICATION NUMBER: 60/090444
67	PRIOR FILING DATE: 1998-06-24
68	PRIOR APPLICATION NUMBER: 60/090445
69	PRIOR FILING DATE: 1998-06-24
70	PRIOR APPLICATION NUMBER: 60/090472
71	PRIOR FILING DATE: 1998-06-24
72	PRIOR APPLICATION NUMBER: 60/090535
73	PRIOR FILING DATE: 1998-06-24
74	PRIOR APPLICATION NUMBER: 60/090540
75	PRIOR FILING DATE: 1998-06-24

1	PRIOR APPLICATION NUMBER: 60/090542
2	PRIOR FILING DATE: 1998-06-24
3	PRIOR APPLICATION NUMBER: 60/090557
4	PRIOR FILING DATE: 1998-06-24
5	PRIOR APPLICATION NUMBER: 60/090676
6	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090678
8	PRIOR FILING DATE: 1998-06-25
9	PRIOR APPLICATION NUMBER: 60/090690
10	PRIOR FILING DATE: 1998-06-25
11	PRIOR APPLICATION NUMBER: 60/090694
12	PRIOR FILING DATE: 1998-06-25
13	PRIOR APPLICATION NUMBER: 60/090695
14	PRIOR FILING DATE: 1998-06-25
15	PRIOR APPLICATION NUMBER: 60/090696
16	PRIOR FILING DATE: 1998-06-25
17	PRIOR APPLICATION NUMBER: 60/090862
18	PRIOR FILING DATE: 1998-06-26
19	PRIOR APPLICATION NUMBER: 60/090863
20	PRIOR FILING DATE: 1998-06-26
21	PRIOR APPLICATION NUMBER: 60/091360
22	PRIOR FILING DATE: 1998-07-01
23	PRIOR APPLICATION NUMBER: 60/091478
24	PRIOR FILING DATE: 1998-07-02
25	PRIOR APPLICATION NUMBER: 60/091544
26	PRIOR FILING DATE: 1998-07-01
27	PRIOR APPLICATION NUMBER: 60/091519
28	PRIOR FILING DATE: 1998-07-02
29	PRIOR APPLICATION NUMBER: 60/091626
30	PRIOR FILING DATE: 1998-07-02
31	PRIOR APPLICATION NUMBER: 60/091633
32	PRIOR FILING DATE: 1998-07-02
33	PRIOR APPLICATION NUMBER: 60/091978
34	PRIOR FILING DATE: 1998-07-07
35	PRIOR APPLICATION NUMBER: 60/091982
36	PRIOR FILING DATE: 1998-07-07
37	PRIOR APPLICATION NUMBER: 60/092182
38	PRIOR FILING DATE: 1998-07-09

Alignment Scores:	1.9e-05	Length:	762
Pred. No.:		Matches:	24
Score:	112.00	Conservative:	6
Percent Similarity:	53.5%	Mismatches:	24
Best Local Similarity:	42.86%	Indels:	2
Query Match:	34.25%	Gaps:	2
DB:	9		

09-833799-13B (1-57) x US-09-990-444-344 (1-762)

Qy	1	AlealnlgluProValIysGlylProValSerThrLysProGlySerCysProIleIleLeu	20
Db	62	GCTGTGAAGAGGATTAAGAGGGTATAGAG---AAAGACGGGGTTTGCCCGAGCTGCACAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProPropAasnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTTCAGTCCGATCTCTCCC---CACTGTTCACACACCAGGACTGCTCTGGGG	175
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	176	GAAGAAGATGTTGTTACCTGTCGACTGTGGCTTCAAGTGTGTGATTCTCT	223

RESULT 12
US-09-989-730-344
; Sequence 344, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC69
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/082250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1,9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 5
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13b (1-57) x US-09-989-730-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAAGAGTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCACAGCTGACAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGATCTCC---CAGTGTACACAGACAGAGCTGTCTGGG 175
QY 41 IleLysLysCysGlyGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 13

US-09-990-436-344
; Sequence 344, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.

;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2730PIC14
;; CURRENT APPLICATION NUMBER: US/09/990,436
;; CURRENT FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167

;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-990-436-344 (1-762)

Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu	20
Db	62	GCTGTGAAGGAGTTAAGAGGGTATAGAG---AAACAGGGGTTTCCCGACTGACAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTTCAAGTCCGATCTCTCCC---CAGTGTACACAGACAGGACTGTCTGGG	175
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	176	GAAGGAAGTGTGTGTACCTGCACCTGTGGTTCAGGTGTGTGATTCTCT	223

RESULT 14

US-09-991-181-344
; Sequence 344, Application US/09991181
; Publication No. US20020197615A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PICS3
; CURRENT APPLICATION NUMBER: US/09/991,181
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23

;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1 9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: Gaps: 2

09-833799-13B (1-57) x US-09-991-181-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTTTCCCGAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 119 GTACGCTGCTTCAAGTCGATCCTCC---CAGTGTCAACAGACCAGGACTGTGTGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACCTGTGCTTCAAGTGTGTGATTCCT 223
RESULT 15
US-09-993-687-344
; Sequence 344, Application US/09993687
; Publication No. US20020198149A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C11
; CURRENT APPLICATION NUMBER: US/09/993.687
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026

1 PRIOR FILING DATE: 1998-06-04
2 PRIOR APPLICATION NUMBER: 60/088028
3 PRIOR FILING DATE: 1998-06-04
4 PRIOR APPLICATION NUMBER: 60/088029
5 PRIOR FILING DATE: 1998-06-04
6 PRIOR APPLICATION NUMBER: 60/088030
7 PRIOR FILING DATE: 1998-06-04
8 PRIOR APPLICATION NUMBER: 60/088033
9 PRIOR FILING DATE: 1998-06-04
10 PRIOR APPLICATION NUMBER: 60/088326
11 PRIOR FILING DATE: 1998-06-04
12 PRIOR APPLICATION NUMBER: 60/088167
13 PRIOR FILING DATE: 1998-06-05
14 PRIOR APPLICATION NUMBER: 60/088202
15 PRIOR FILING DATE: 1998-06-05
16 PRIOR APPLICATION NUMBER: 60/088212
17 PRIOR FILING DATE: 1998-06-05
18 PRIOR APPLICATION NUMBER: 60/088217
19 PRIOR FILING DATE: 1998-06-05
20 PRIOR APPLICATION NUMBER: 60/088655
21 PRIOR FILING DATE: 1998-06-09
22 PRIOR APPLICATION NUMBER: 60/088734
23 PRIOR FILING DATE: 1998-06-10
24 PRIOR APPLICATION NUMBER: 60/088738
25 PRIOR FILING DATE: 1998-06-10
26 PRIOR APPLICATION NUMBER: 60/088742
27 PRIOR FILING DATE: 1998-06-10
28 PRIOR APPLICATION NUMBER: 60/088810
29 PRIOR FILING DATE: 1998-06-10
30 PRIOR APPLICATION NUMBER: 60/088824
31 PRIOR FILING DATE: 1998-06-10
32 PRIOR APPLICATION NUMBER: 60/088826
33 PRIOR FILING DATE: 1998-06-10
34 PRIOR APPLICATION NUMBER: 60/088858
35 PRIOR FILING DATE: 1998-06-11
36 PRIOR APPLICATION NUMBER: 60/088861
37 PRIOR FILING DATE: 1998-06-11
38 PRIOR APPLICATION NUMBER: 60/088876
39 PRIOR FILING DATE: 1998-06-11
40 PRIOR APPLICATION NUMBER: 60/089105
41 PRIOR FILING DATE: 1998-06-12
42 PRIOR APPLICATION NUMBER: 60/089440
43 PRIOR FILING DATE: 1998-06-16
44 PRIOR APPLICATION NUMBER: 60/089512
45 PRIOR FILING DATE: 1998-06-16
46 PRIOR APPLICATION NUMBER: 60/089514
47 PRIOR FILING DATE: 1998-06-16
48 PRIOR APPLICATION NUMBER: 60/089532
49 PRIOR FILING DATE: 1998-06-17
50 PRIOR APPLICATION NUMBER: 60/089538
51 PRIOR FILING DATE: 1998-06-17
52 PRIOR APPLICATION NUMBER: 60/089598
53 PRIOR FILING DATE: 1998-06-17
54 PRIOR APPLICATION NUMBER: 60/089599
55 PRIOR FILING DATE: 1998-06-17
56 PRIOR APPLICATION NUMBER: 60/089600
57 PRIOR FILING DATE: 1998-06-17
58 PRIOR APPLICATION NUMBER: 60/089653
59 PRIOR FILING DATE: 1998-06-17
60 PRIOR APPLICATION NUMBER: 60/089801
61 PRIOR FILING DATE: 1998-06-18
62 PRIOR APPLICATION NUMBER: 60/089907
63 PRIOR FILING DATE: 1998-06-18
64 PRIOR APPLICATION NUMBER: 60/089908
65 PRIOR FILING DATE: 1998-06-18
66 PRIOR APPLICATION NUMBER: 60/089947
67 PRIOR FILING DATE: 1998-06-19
68 PRIOR APPLICATION NUMBER: 60/089948
69 PRIOR FILING DATE: 1998-06-19
70 PRIOR APPLICATION NUMBER: 60/089952
71 PRIOR FILING DATE: 1998-06-19
72 PRIOR APPLICATION NUMBER: 60/090246
73 PRIOR FILING DATE: 1998-06-22

74 PRIOR APPLICATION NUMBER: 60/090252
75 PRIOR FILING DATE: 1998-06-22
76 PRIOR APPLICATION NUMBER: 60/090254
77 PRIOR FILING DATE: 1998-06-22
78 PRIOR APPLICATION NUMBER: 60/090349
79 PRIOR FILING DATE: 1998-06-23
80 PRIOR APPLICATION NUMBER: 60/090355
81 PRIOR FILING DATE: 1998-06-23
82 PRIOR APPLICATION NUMBER: 60/090429
83 PRIOR FILING DATE: 1998-06-24
84 PRIOR APPLICATION NUMBER: 60/090431
85 PRIOR FILING DATE: 1998-06-24
86 PRIOR APPLICATION NUMBER: 60/090435
87 PRIOR FILING DATE: 1998-06-24
88 PRIOR APPLICATION NUMBER: 60/090444
89 PRIOR FILING DATE: 1998-06-24
90 PRIOR APPLICATION NUMBER: 60/090445
91 PRIOR FILING DATE: 1998-06-24
92 PRIOR APPLICATION NUMBER: 60/090472
93 PRIOR FILING DATE: 1998-06-24
94 PRIOR APPLICATION NUMBER: 60/090535
95 PRIOR FILING DATE: 1998-06-24
96 PRIOR APPLICATION NUMBER: 60/090540
97 PRIOR FILING DATE: 1998-06-24
98 PRIOR APPLICATION NUMBER: 60/090542
99 PRIOR FILING DATE: 1998-06-24
100 PRIOR APPLICATION NUMBER: 60/090557
101 PRIOR FILING DATE: 1998-06-24
102 PRIOR APPLICATION NUMBER: 60/090676
103 PRIOR FILING DATE: 1998-06-25
104 PRIOR APPLICATION NUMBER: 60/090678
105 PRIOR FILING DATE: 1998-06-25
106 PRIOR APPLICATION NUMBER: 60/090690
107 PRIOR FILING DATE: 1998-06-25
108 PRIOR APPLICATION NUMBER: 60/090694
109 PRIOR FILING DATE: 1998-06-25
110 PRIOR APPLICATION NUMBER: 60/090695
111 PRIOR FILING DATE: 1998-06-25
112 PRIOR APPLICATION NUMBER: 60/090696
113 PRIOR FILING DATE: 1998-06-25
114 PRIOR APPLICATION NUMBER: 60/090862
115 PRIOR FILING DATE: 1998-06-26
116 PRIOR APPLICATION NUMBER: 60/090863
117 PRIOR FILING DATE: 1998-06-26
118 PRIOR APPLICATION NUMBER: 60/091360
119 PRIOR FILING DATE: 1998-07-01
120 PRIOR APPLICATION NUMBER: 60/091478
121 PRIOR FILING DATE: 1998-07-02
122 PRIOR APPLICATION NUMBER: 60/091544
123 PRIOR FILING DATE: 1998-07-01
124 PRIOR APPLICATION NUMBER: 60/091519
125 PRIOR FILING DATE: 1998-07-02
126 PRIOR APPLICATION NUMBER: 60/091626
127 PRIOR FILING DATE: 1998-07-02
128 PRIOR APPLICATION NUMBER: 60/091633
129 PRIOR FILING DATE: 1998-07-02
130 PRIOR APPLICATION NUMBER: 60/091978
131 PRIOR FILING DATE: 1998-07-07
132 PRIOR APPLICATION NUMBER: 60/091982
133 PRIOR FILING DATE: 1998-07-07
134 PRIOR APPLICATION NUMBER: 60/092182
135 PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-993-687-344 (1-762)


```

Qy 1 AlaGlnGluProValysGlyProValserThrIysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAAGGAGTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCCAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGCGATCCTCCC---CAGTGTCCACACACACAGGAGCTGTCTGGGG 175
Qy 41 IleIysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 16
US-09-989-734-344
; Sequence 344, Application US/09989734
; Publication No. US20030003531A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC64
; CURRENT APPLICATION NUMBER: US/09/989,734
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827

```

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-989-734-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAGAGGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTGCCACGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGAGCTCTCCC---CAGTGTACACAGCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAGGTGTGTACTGCTGCTTCAAGTGTGATTCTTCT 223

RESULT 17

US-09-997-653-344
; Sequence 344, Application US/09997653
; Publication No. US20030008297A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC38
; CURRENT APPLICATION NUMBER: US/09/997,653
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05
Score: 112.00
Percent Similarity: 53.57%

Length: 762
Matches: 24
Conservative: 6

[illegible]

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

```

; PRIOR FILING DATE: 1998-07-09
Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Conservative: 6
Percent Similarity: 53.57% Mismatches: 24
Best Local Similarity: 42.86% Indels: 2
Query Match: 34.25% Gaps: 2
DB:

09-833799-13B (1-57) x US-09-997-653-344 (1-762)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
DB 62 GCTGTGGAAGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
DB 119 GTACGCTGTTCAAGTCCGATCTCTCC---CAGTGTCAACAGACAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
DB 176 GAAAGGAGTGTGTGTACCTGCCTGCTGCTTCAAGTGTGTGATCTCT 223

RESULT 18
US-09-993-667-344
; Sequence 344, Application US/09993667
; Publication No. US20030022187A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gernitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC4
; CURRENT APPLICATION NUMBER: US/09/993,667
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17

```

; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633

; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-993-667-344 (1-762)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGTTTGCACGCTGACAAAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGCGATCCTCCC---CAGTGTCAACACAGACCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCTCT 223

RESULT 19

US-09-990-438-344
; Sequence 344, Application US/09990438
; Publication No. US20030027754A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC3
; CURRENT APPLICATION NUMBER: US/09/990,438
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1,9e-05 Length: 762
 Score: 112.00 Matches: 24
 Percent Similarity: 53.57% Conservative: 6
 Best Local Similarity: 42.86% Mismatches: 24
 Query Match: 34.25% Indels: 2
 DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-990-438-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 62 GCTGTGGAAGGAGTTAAAGAGGTATAGAG---AAAGCAGGGGTTTCCCCAGCTGACAAC 118
 Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 119 GTACGCTGCTCAAGTCGATCCTCC---CAGTGTACACAGACAGGACTGCTGGG 175
 Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 176 GAAAGGAAGTGTGTACTCCTGCTTCAAGTGTGATTCCT 223

RESULT 20

US-09-990-562-344

; Sequence 344, Application US/09990562

; Publication No. US20030027985A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas P.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2730P1C18
 ; CURRENT APPLICATION NUMBER: US/09/990,562
 ; CURRENT FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16

;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863

;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1,9e-05
Score: 112.00
Percent Similarity: 53.57%
Best Local Similarity: 42.86%
Query Match: 34.25%
DB: 9

Length: 762
Matches: 24
Conservative: 6
Mismatches: 24
Indels: 2
Gaps: 2

09-833799-13B (1-57) x US-09-990-562-344 (1-762)

QY 1 AlaGlnGluProValysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGTCTCAAGTCCGATCTCTCC---CAGTGTACACAGACAGAGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 21
US-09-997-428-344
; Sequence 344, Application US/09997428
; Publication No. US20030027162A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

[illegible]

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

```

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C42
; CURRENT FILING DATE: 2001-11-15
; CURRENT APPLICATION NUMBER: US/09/997,666
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-997-428-344 (1-762)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGCACGACTGACAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGTTCAAGTCGATCCCTCC---CAGTGTACACAGACCCAGGACTGTCTGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTTACCTGTCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 22
US-09-997-666-344
; Sequence 344, Application US/09997666
; Publication No. US20030027163A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deshoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
```

1	PRIOR APPLICATION NUMBER: 60/0888262
2	PRIOR FILING DATE: 1998-06-10
3	PRIOR APPLICATION NUMBER: 60/088858
4	PRIOR FILING DATE: 1998-06-11
5	PRIOR APPLICATION NUMBER: 60/088861
6	PRIOR FILING DATE: 1998-06-11
7	PRIOR APPLICATION NUMBER: 60/088876
8	PRIOR FILING DATE: 1998-06-11
9	PRIOR APPLICATION NUMBER: 60/089105
10	PRIOR FILING DATE: 1998-06-12
11	PRIOR APPLICATION NUMBER: 60/089440
12	PRIOR FILING DATE: 1998-06-16
13	PRIOR APPLICATION NUMBER: 60/089512
14	PRIOR FILING DATE: 1998-06-16
15	PRIOR APPLICATION NUMBER: 60/089514
16	PRIOR FILING DATE: 1998-06-16
17	PRIOR APPLICATION NUMBER: 60/089532
18	PRIOR FILING DATE: 1998-06-17
19	PRIOR APPLICATION NUMBER: 60/089533
20	PRIOR FILING DATE: 1998-06-17
21	PRIOR APPLICATION NUMBER: 60/089598
22	PRIOR FILING DATE: 1998-06-17
23	PRIOR APPLICATION NUMBER: 60/089599
24	PRIOR FILING DATE: 1998-06-17
25	PRIOR APPLICATION NUMBER: 60/089600
26	PRIOR FILING DATE: 1998-06-17
27	PRIOR APPLICATION NUMBER: 60/089653
28	PRIOR FILING DATE: 1998-06-17
29	PRIOR APPLICATION NUMBER: 60/089801
30	PRIOR FILING DATE: 1998-06-18
31	PRIOR APPLICATION NUMBER: 60/089907
32	PRIOR FILING DATE: 1998-06-18
33	PRIOR APPLICATION NUMBER: 60/089908
34	PRIOR FILING DATE: 1998-06-18
35	PRIOR APPLICATION NUMBER: 60/089947
36	PRIOR FILING DATE: 1998-06-19
37	PRIOR APPLICATION NUMBER: 60/089948
38	PRIOR FILING DATE: 1998-06-19
39	PRIOR APPLICATION NUMBER: 60/089952
40	PRIOR FILING DATE: 1998-06-19
41	PRIOR APPLICATION NUMBER: 60/090246
42	PRIOR FILING DATE: 1998-06-22
43	PRIOR APPLICATION NUMBER: 60/090252
44	PRIOR FILING DATE: 1998-06-22
45	PRIOR APPLICATION NUMBER: 60/090254
46	PRIOR FILING DATE: 1998-06-22
47	PRIOR APPLICATION NUMBER: 60/090349
48	PRIOR FILING DATE: 1998-06-23
49	PRIOR APPLICATION NUMBER: 60/090355
50	PRIOR FILING DATE: 1998-06-23
51	PRIOR APPLICATION NUMBER: 60/090429
52	PRIOR FILING DATE: 1998-06-24
53	PRIOR APPLICATION NUMBER: 60/090431
54	PRIOR FILING DATE: 1998-06-24
55	PRIOR APPLICATION NUMBER: 60/090435
56	PRIOR FILING DATE: 1998-06-24
57	PRIOR APPLICATION NUMBER: 60/090444
58	PRIOR FILING DATE: 1998-06-24
59	PRIOR APPLICATION NUMBER: 60/090445
60	PRIOR FILING DATE: 1998-06-24
61	PRIOR APPLICATION NUMBER: 60/090472
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/090535
64	PRIOR FILING DATE: 1998-06-24
65	PRIOR APPLICATION NUMBER: 60/090540
66	PRIOR FILING DATE: 1998-06-24
67	PRIOR APPLICATION NUMBER: 60/090542
68	PRIOR FILING DATE: 1998-06-24
69	PRIOR APPLICATION NUMBER: 60/090557
70	PRIOR FILING DATE: 1998-06-24
71	PRIOR APPLICATION NUMBER: 60/090676
72	PRIOR FILING DATE: 1998-06-25
73	PRIOR APPLICATION NUMBER: 60/090678
74	PRIOR FILING DATE: 1998-06-25

7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090690	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090694	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090695	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090696	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090862	
7	PRIOR FILING DATE: 1998-06-26	
7	PRIOR APPLICATION NUMBER: 60/090863	
7	PRIOR FILING DATE: 1998-06-26	
7	PRIOR APPLICATION NUMBER: 60/091360	
7	PRIOR FILING DATE: 1998-07-01	
7	PRIOR APPLICATION NUMBER: 60/091478	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091544	
7	PRIOR FILING DATE: 1998-07-01	
7	PRIOR APPLICATION NUMBER: 60/091519	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091626	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091633	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091978	
7	PRIOR FILING DATE: 1998-07-07	
7	PRIOR APPLICATION NUMBER: 60/091982	
7	PRIOR FILING DATE: 1998-07-07	
7	PRIOR APPLICATION NUMBER: 60/092182	
7	PRIOR FILING DATE: 1998-07-07	

Alignment Scores:

Alignment Scores:			
Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-997-666-344 (1-762)

Qy	1	AlaGlnGluProValLysGlyProValSerThrylsProGlysSerCysProIleLeu	20
Db	62	GCTGTGAAGGAGTTAAAGAGGGTATAG-- - AAAGCAGGGGTTTGCCAGCTGACAAC	118
		: : :	
Qy	21	IleArgcysAlaMetLeuAsnProProIsnArgCysLeuLysAspThrAscCysProGly	40
		: : : : : : : :	
Db	119	GTACGCGTCTCAAGTCGCATCCTCCC- - CAGTGTCAACAGACCAGGACTGTCTGGGG	175
		: : :	
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
		: : : : : : : :	
Db	176	GAAGAAGTGTTGTTACTGCACTGTGGCTTCAAGTGTGTGATTCCT	223
		: : :	

RESULT 23

US-09-989-722-344
RESULTS 23

03-03-363-722-344
; Sequence 344, Application US/09989722

; Patent No. US20020072067A1

; GENERAL INFORMATION:

APPLICANT:	Ashtenazi, Avi J.
APPLICANT:	Baker, Kevin P.
APPLICANT:	Botstein, David
APPLICANT:	Desnoyers, Luc
APPLICANT:	Eaton, Dan L.
APPLICANT:	Ferrara, Napoleone
APPLICANT:	Fong, Sherman
APPLICANT:	Gerber, Hanspeter
APPLICANT:	Gerritsen, Mary E.
APPLICANT:	Goddard, Audrey
APPLICANT:	Godowski, Paul J.
APPLICANT:	Grimaldi, J. Christopher
APPLICANT:	Gurney, Austin L.
APPLICANT:	Klavin, Ivar J.

APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PLC63
CURRENT APPLICATION NUMBER: US/09/989,722
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-989-722-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTTGCCCGAGCTGACAAAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCTCCC---CAGTGTCACACAGCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysGlyGluGlySerGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCCTGTGCTTCAAGTGTGTGATTCCT 223

RESULT 24

US-09-989-723-344
; Sequence 344, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter

;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; TITLE OF INVENTION: Acids Encoding the Same
;; FILE REFERENCE: P2730PIC62
;; CURRENT APPLICATION NUMBER: US/09/989,723
;; CURRENT FILING DATE: 2001-11-19
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05

;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05
Score: 112.00
Percent Similarity: 53.57%
Best Local Similarity: 42.86%
Query Match: 34.25%
DB: 10
Length: 762
Matches: 24
Conservative: 6
Mismatch: 24
Indels: 2
Gaps: 2

09-833799-13B (1-57) x US-09-989-723-344 (1-762)
Qy 1 AlaGlnProValIysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGGTGACAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCTCTCCC---CAGTGTACACACAGACAGGACTGTCTGGG 175
Qy 41 IleIysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTTACCTGCACTGTGGCTTCAAGTGTGTGATTCT 223
RESULT 25
US-09-989-279-344
; Sequence 344, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David
APPLICANT: Denoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C56
CURRENT APPLICATION NUMBER: US/09/989,279
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.5%	Conservative:	6
Best Local Similarity:	42.8%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-989-279-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGAGGTAAAGAGGGTATAGAG---AAAGCAGGGGTTGGCCCACTGACAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCC---CAGTGTCAACAGACAGGACTGTCTGGG 175
QY 41 IleLysLysCysGlyGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACCTGACTGTGGCTTCAAGTGTGTGATTCCT 223
RESULT 26

US-09-989-727-344
; Sequence 344, Application US/09989727
; Patent No. US20020072497A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355

;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-989-727-344 (1-762)

QY	1	AlaGlnGluProValIysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	62	GCTGTGGAAGAGATTAAAGAGGGTATAGAG---	AAAGAGGGGTTTGCCAGCTGACAC 118
QY	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrCysProGly	40
Db	119	GTACGCTGCTCAAGTCCGATCCTCCC---	CAGTGTACACAGACGAGACTGTCTGGGG 175

Qy 41 IleLysLysCysCysGluClySerCysGlyMetalCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACTGCACCTGTGCTTCAAGTGTGTGATTCCT 223

RESULT 27
US-09-989-731-344
; Sequence 344. Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC70
; CURRENT APPLICATION NUMBER: US/09/989,731
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252

;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-989-731-344 (1-762)

QY 1 AlaGlnCluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCCACGCTGACAAAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTCAACACAGACGAGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCTCT 223

RESULT 28

US-09-989-732-344
; Sequence 344, Application US/09989732
; Patent No. US20020123463A1

GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi J.
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C57
; CURRENT APPLICATION NUMBER: US/09/989,732
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03

Alignment Scores:	1.9e-05	762
Pred. No.:		
Score:	115.00	24
Percent Similarity:	53.57%	6
Best Local Similarity:	42.86%	24
Mismatches:		
Matches:		
Length:		

Alignment Scores:	
Pred. No.:	1.9e-05
Score:	112.00
Percent Similarity:	53.57%
Best Local Similarity:	42.86%

```
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2
09-833799-13B (1-57) x US-09-989-732-344 (1-762)
QY 1 AlaGlnGluProValysGlyProValserThrlyserProGlyserCysProIlelleLeu 20
Db 62 GCTGTGAAGAGGTTAAAGAGGGTATAGAG---AAACGAGGGTTTCCCGAGCTGACAC 118
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuIlysAspThrAspCysProGly 40
Db 119 GTACGCTGCTCAAGTCCGATCCTCCC---CAGTGTACACACAGCAGGACTGCTGGGG 175
QY 41 IleIyslysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACGCTGTGGTTCAGAGTGTGTGATTCT 223
RESULT 29
US-09-991-073-344
; Sequence 344, Application US/09991073
; Patent No. US20020127576A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC15
; CURRENT APPLICATION NUMBER: US/09/991,073
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
```

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Alignment Scores: 1.9e-05 Length: 762
 Pred. No.: 112.00 Matches: 24
 Score: 53.57% Conservative: 24
 Percent Similarity: 42.86% Mismatches: 2
 Best Local Similarity: 34.25% Indels: 2
 Query Match: 10 Gaps: 2
 DB:

09-833799-13B (1-57) x US-09-991-073-344 (1-762)
 QY 1 AlaGlnGluProValysGlyProVaiserThrLysProGlySerCysProIleIleIeu 20
 Db 62 GCTGTGGAGGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAAC 118
 QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 119 GTACGCTGCTTCAAGTCGATCCTCCC---CAGTGTACACAGACAGGAGCTGTCTGGG 175
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 176 GAAGGAAGAGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 30
 US-09-990-442-344
 ; Sequence 344, Application US/09990442
 ; Patent No. US20020132252A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Geritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kijavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2730PIC8
 ; CURRENT APPLICATION NUMBER: US/09/990,442
 ; CURRENT FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28

1	PRIOR APPLICATION NUMBER: 60/084600
2	PRIOR FILING DATE: 1998-05-07
3	PRIOR APPLICATION NUMBER: 60/087106
4	PRIOR FILING DATE: 1998-05-28
5	PRIOR APPLICATION NUMBER: 60/087607
6	PRIOR FILING DATE: 1998-06-02
7	PRIOR APPLICATION NUMBER: 60/087609
8	PRIOR FILING DATE: 1998-06-02
9	PRIOR APPLICATION NUMBER: 60/087759
10	PRIOR FILING DATE: 1998-06-02
11	PRIOR APPLICATION NUMBER: 60/087827
12	PRIOR FILING DATE: 1998-06-03
13	PRIOR APPLICATION NUMBER: 60/088021
14	PRIOR FILING DATE: 1998-06-04
15	PRIOR APPLICATION NUMBER: 60/088025
16	PRIOR FILING DATE: 1998-06-04
17	PRIOR APPLICATION NUMBER: 60/088026
18	PRIOR FILING DATE: 1998-06-04
19	PRIOR APPLICATION NUMBER: 60/088028
20	PRIOR FILING DATE: 1998-06-04
21	PRIOR APPLICATION NUMBER: 60/088029
22	PRIOR FILING DATE: 1998-06-04
23	PRIOR APPLICATION NUMBER: 60/088326
24	PRIOR FILING DATE: 1998-06-04
25	PRIOR APPLICATION NUMBER: 60/088167
26	PRIOR FILING DATE: 1998-06-05
27	PRIOR APPLICATION NUMBER: 60/088202
28	PRIOR FILING DATE: 1998-06-05
29	PRIOR APPLICATION NUMBER: 60/088212
30	PRIOR FILING DATE: 1998-06-05
31	PRIOR APPLICATION NUMBER: 60/088217
32	PRIOR FILING DATE: 1998-06-05
33	PRIOR APPLICATION NUMBER: 60/088655
34	PRIOR FILING DATE: 1998-06-09
35	PRIOR APPLICATION NUMBER: 60/088734
36	PRIOR FILING DATE: 1998-06-10
37	PRIOR APPLICATION NUMBER: 60/088738
38	PRIOR FILING DATE: 1998-06-10
39	PRIOR APPLICATION NUMBER: 60/088742
40	PRIOR FILING DATE: 1998-06-10
41	PRIOR APPLICATION NUMBER: 60/088810
42	PRIOR FILING DATE: 1998-06-10
43	PRIOR APPLICATION NUMBER: 60/088824
44	PRIOR FILING DATE: 1998-06-10
45	PRIOR APPLICATION NUMBER: 60/088826
46	PRIOR FILING DATE: 1998-06-10
47	PRIOR APPLICATION NUMBER: 60/088858
48	PRIOR FILING DATE: 1998-06-11
49	PRIOR APPLICATION NUMBER: 60/088861
50	PRIOR FILING DATE: 1998-06-11
51	PRIOR APPLICATION NUMBER: 60/088876
52	PRIOR FILING DATE: 1998-06-11
53	PRIOR APPLICATION NUMBER: 60/089105
54	PRIOR FILING DATE: 1998-06-12
55	PRIOR APPLICATION NUMBER: 60/089532
56	PRIOR FILING DATE: 1998-06-17
57	PRIOR APPLICATION NUMBER: 60/089538
58	PRIOR FILING DATE: 1998-06-17
59	PRIOR APPLICATION NUMBER: 60/089598
60	PRIOR FILING DATE: 1998-06-17
61	PRIOR APPLICATION NUMBER: 60/089599
62	PRIOR FILING DATE: 1998-06-17
63	PRIOR APPLICATION NUMBER: 60/089600

Sun Feb 16 09:13:26 2003

09-833799-13b.rnpb

```
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.:      1,98-05      Length:      762
Score:          112.00      Matches:      24
Percent Similarity: 53.57%      Conservative: 6
Best Local Similarity: 42.86%      Mismatches: 24
Query Match:      34.25%      Indels:      2
DB:              10          Gaps:      2

09-833799-13B (1-57) x US-09-990-442-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGAGGATTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCACGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTCAAGTCGATCTCCC---CAGTGTCAACAGACCAGGACTGCTGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAGTGTGTACCTGCACGTGCTGCTTCAAGTGTGTGATTCCT 223

RESULT 31
US-09-991-163-344
; Sequence 344, Application US/09991163
; Patent No. US20020132253A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C17
; CURRENT APPLICATION NUMBER: US/09/991,163
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
```


; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01

; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-991-163-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTCCACAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAspProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACCGTCTTCAAGTCGAGCTCCCTCCC---CAGTGTACACAGACCAGAGCTGTCTGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 32

US-09-993-604-344
; Sequence 344, Application US/09993604
; Patent No. US20020137075A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: KJavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

09-833799-13b.rnpb

```
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05          Length: 762
Score: 112.00             Matches: 24
Percent Similarity: 53.57%  Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25%        Indels: 2
DB: 10                     Gaps: 2

09-833799-13B (1-57) x US-09-993-604-344 (1-762)

Qy 1 AlaGInGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGAGGATTAAGAGGGTATAGAG---AAAGCAGGGGTTTCCCGAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAGTCCGATCTCTCC---CAGTGTACACAGACAGAGACTGTCTGGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTCTTACTCCTGCATGCTGGCTTCAAGTGTGTGATTCCT 223

RESULT 33
US-09-990-456-344
; Sequence 344, Application US/09990456
; Patent No. US20020137890A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C22

; CURRENT APPLICATION NUMBER: US/09/990,456
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
```

; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1-9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-990-456-344 (1-762)

Qy 1 AlaGlnGluProValIysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCTCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACACTGTGGCTTCAAGTGTGATTCCT 223

RESULT 34

US-09-989-721-344
; Sequence 344, Application US/09989721
; Patent No. US20020142961A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Pong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P27301C55
CURRENT APPLICATION NUMBER: US/09/989,721
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-989-721-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProLeuLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGGCCAGCTGACAAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCC---CAGTGTACACACAGCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTTGTACCTGCACCTGTGGCTTCAAGTGTGTATCTCT 223

RESULT 35

US-09-964-824A-495/c
; Sequence 495, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495

; LENGTH: 411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(411)
; OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-495

Alignment Scores:

Pred. No.: 3.58e-05 Length: 411
Score: 107.00 Matches: 21
Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 10 Gaps: 0

09-833799-13B (1-57) x US-09-964-824A-495 (1-411)

QY 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCACACAGGNGNAACCTTGGNAAGTNCNAGTGANTTNGCCCAATGTTNGATGCTNAA 282
QY 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyLleLysLysCysCysGluGl 47
Db 281 CCCCCCAATTTCTGTGAGATGGATGGCCAGTCGCAAGCGTGACTTGAAGTGTTCATGGG 222
QY 47 YSerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGAAATCTCGCTTTCCTTCCCT 194

RESULT 36

US-09-954-456-1987/c
; Sequence 1987, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1987
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1987

Alignment Scores:

Pred. No.: 3,58e-05 Length: 411
Score: 107.00 Matches: 21
Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 10 Gaps: 0

09-833799-13B (1-57) x US-09-954-456-1987 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAGGNGNAACCTTGGNAAGTCCNAGTGANTTTGGCCCAATGTTNGATGCTNAA 282
Qy 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysGluG1 47
Db 281 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAGCGTGACTTGAAGTGTTCATGGG 222
Qy 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGAAATCCTCGCTTCCCT 194

RESULT 37

US-09-967-768A-3/C

; Sequence 3, Application US/09967768A

; Patent No. US20020150877A1

; GENERAL INFORMATION:

; APPLICANT: Augustus, Meena

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; FILE REFERENCE: 689290-72

; CURRENT APPLICATION NUMBER: US/09/967,768A

; CURRENT FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US/60/236,109

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,034

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,111

; PRIOR FILING DATE: 2000-09-28

; NUMBER OF SEQ ID NOS: 325

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 411

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)...(411)

; OTHER INFORMATION: n=a,t,g or c

US-09-967-768A-3

Alignment Scores:

Pred. No.: 3,58e-05 Length: 411
Score: 107.00 Matches: 21
Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 10 Gaps: 0

09-833799-13B (1-57) x US-09-967-768A-3 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAGGNGNAACCTTGGNAAGTCCNAGTGANTTTGGCCCAATGTTNGATGCTNAA 282
Qy 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysGluG1 47
Db 281 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAGCGTGACTTGAAGTGTTCATGGG 222

Qy 47 ySerCysGlyMetAlaCysPheValPro 56

Db 221 CATGTGTGGAAATCCTCGCTTCCCT 194

RESULT 38

US-09-884-441-377
; Sequence 377, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:

; APPLICANT: Algater, Paul A.

; APPLICANT: Carter, Darrick

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.462C7

; CURRENT APPLICATION NUMBER: US/09/884,441

; CURRENT FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 489

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 377

; LENGTH: 292

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-884-441-377

Alignment Scores:

Pred. No.: 4,07e-05 Length: 292
Score: 105.00 Matches: 18
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 51.43% Mismatches: 15
Query Match: 32.11% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x US-09-884-441-377 (1-292)

Qy 22 ArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1 CAATGTTTGTGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCACGCGTGAC 60

Qy 42 LysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 61 TTGAAGTGTTCATGGGATGTGTGGAAATCCTCGCTTCCCT 105

RESULT 39

US-10-152-661-589

; Sequence 589, Application US/10152661

; Publication No. US20030022835A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; FILE REFERENCE: 11000.1011c5

; CURRENT APPLICATION NUMBER: US/10/152,661

; CURRENT FILING DATE: 2002-05-20

; PRIOR APPLICATION NUMBER: 09/866,050

; PRIOR FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: 60/221,232

; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: 60/206,650

; PRIOR FILING DATE: 2000-05-24

; PRIOR APPLICATION NUMBER: 09/312,283

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: PCT/NZ99/00051

; PRIOR FILING DATE: 1999-04-29

; PRIOR APPLICATION NUMBER: 09/188,930

; PRIOR FILING DATE: 1998-11-09

; PRIOR APPLICATION NUMBER: 09/069,726

; PRIOR FILING DATE: 1998-04-29

; NUMBER OF SEQ ID NOS: 725

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 589

; LENGTH: 724

; TYPE: DNA

; ORGANISM: Mouse
US-10-152-661-589

Alignment Scores:

Pred. No.: 0.00103 Length: 724
Score: 98.00 Matches: 25
Percent Similarity: 43.24% Conservative: 7
Best Local Similarity: 33.78% Mismatches: 18
Query Match: 29.97% Indels: 24
DB: 9 Gaps: 4

09-833799-13B (1-57) x US-10-152-661-589 (1-724)

QY 3 GluProValLysGlyProValSerThrLysPro----- 13

Db 313 AAACCCAGGAGGTCAAGTCTCCAGAACCCAGCGGTGTGACCAAGGAGGCTTAGGT 372

QY 14 -----GlySerCysProLlelle-----LeulleArgCysala 24

Db 373 GTCCGAGAAAAGCAGGACACCTTCCCGCCAGCGTGGACATACCCAAAGCTCGGCTCTGT 429

QY 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyLysLysCys 44

Db 430 -----GAGCACCAGTGTGAGTGGACAGCCAGTGTCTGGCAACATGAATGC 477

QY 45 CysGluGlySerCysGly---MetAlaCysPheValProGln 57

Db 478 TGCCGCAATGGATGTGGGAAGATGGCTGCACACACCCAAA 519

RESULT 40

US-09-967-768A-214

; Sequence 214, Application US/09967768A

; Patent No. US20020150877A1

; GENERAL INFORMATION:

; APPLICANT: Augustus, Meena

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; FILE REFERENCE: 689290-72

; CURRENT APPLICATION NUMBER: US/09/967,768A

; PRIOR FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US/60/236,109

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,034

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,111

; NUMBER OF SEQ ID NOS: 325

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 214

; LENGTH: 583

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-967-768A-214

Alignment Scores:

Pred. No.: 0.00103 Length: 583
Score: 97.00 Matches: 20
Percent Similarity: 53.19% Conservative: 5
Best Local Similarity: 42.55% Mismatches: 20
Query Match: 29.66% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-967-768A-214 (1-583)

QY 12 LysProGlySerCysProLlelleLeulleArgCysAlaMetLeuAsn---ProProAsn 30

Db 256 AAGGAGGTTCTCGCCCGCCAGGTGAACATTAATTTCCCGAGCTCGGCTCTGTGGGAC 315

QY 31 ArgCysLeuLysAspThrAspCysProGlyLleLysCysGluCysSerCysGly 50

Db 316 CATGCCAGGTGACACGAGTGTCTGCCAGCATGAATGTCCGCAATGGCTGTGGG 375

QY 51 ---MetAlaCysPheValPro 56

Db 376 AAGGTGTCTGTGTGTCTCTCC 396

RESULT 41

US-09-864-761-1629

; Sequence 1629, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aeonica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 1629

; LENGTH: 469

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AL050348.19

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5

Alignment Scores:

Pred. No.: 0.00138 Length: 469
Score: 95.00 Matches: 19
Percent Similarity: 47.37% Conservative: 8
Best Local Similarity: 33.33% Mismatches: 22

Query Match: 29.05% Indels: 8
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-864-761-1629 (1-469)

QY 4 ProValLysGlyProValSerThrLysPro-----GlySerCysProIlelle 19
|||||
DB 228 CCTGTCCCTTTCTGTGCGCTGTGTTCCACAGGGCGGGCGGTGATTGTCCAAAAGTT 287
|||||

QY 20 LeuileArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysPro 39
|||||
DB 288 CTGTGGGCGCTGTGCATTGTGGC-----TGTGTGGATGAGAAATTGTCAA 335
|||||

QY 40 GlylleLysLysCysGlyGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
DB 336 GCTGGAGAAAATGTCAGTCAAGTCAGGCTGTGGCGCTTCTGTGTCCACCA 386
|||||

RESULT 42
US-09-790-264-57
; Sequence 57, Application US/09790264
; Patent No. US20020028508A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 07334-322001
; CURRENT APPLICATION NUMBER: US/09/790,264
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/102,705
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: US 09/363,630
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-790-264-57

Alignment Scores:
Pred. No.: 0.000616 Length: 228
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservative: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-790-264-57 (1-228)

QY 12 LysProGlySerCysProIlelleLeuileArgCysAlaMetLeuAsnProPro----- 29
|||||
DB 91 AAACCTGAGCTGTGCTCC-----AAGCCTTCACCAGAA 123
|||||

QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlylleLysLys 43
|||||
DB 124 AGTGTGGAATTTGTGTGATCAATGCTCAGGAGATGATGCTGCTGCGCAACATGAAG 183
|||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||

DB 184 TGCTGTAGCAATAGCTGTGTGTCATGTCGCAAAACTCCT 222

RESULT 43
US-09-823-038A-29
; Sequence 29, Application US/09823038A
; Patent No. US20020058335A1
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1037c3
; CURRENT APPLICATION NUMBER: US/09/823,038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Mouse
US-09-823-038A-29

Alignment Scores:
Pred. No.: 0.00146 Length: 437
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservative: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-823-038A-29 (1-437)

QY 12 LysProGlySerCysProIlelleLeuileArgCysAlaMetLeuAsnProPro----- 29
|||||
DB 80 AAACCTGAGCTGTGCTCC-----AAGCCTTCACCAGAA 112
|||||

QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlylleLysLys 43
|||||
DB 113 AGTGTGGAATTTGTGTGATCAATGCTCAGGAGATGATGCTGCTGCGCAACATGAAG 172
|||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||

DB 173 TGCTGTAGCAATAGCTGTGTGTCATGTCGCAAAACTCCT 211

RESULT 44
US-09-790-264-55
; Sequence 55, Application US/09790264
; Patent No. US20020028508A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 07334-322001
; CURRENT APPLICATION NUMBER: US/09/790,264
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/102,705
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: US 09/363,630

```
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(264)
US-09-790-264-55

Alignment Scores:
Pred. No.: 0.00161 Length: 471
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservative: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-09-790-264-55 (1-471)

Qy 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProPro----- 29
Db 127 AAACCTGGAGCTGTGCC-----AAGCCTTACCAGAA 159
Qy 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 160 AGTGTGGAAATTGTTGATCAATGCTCAGGAGATGATCCTCCCTGGCAACATGAAG 219
Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 220 TGCTGTAGCAATAGCTGTGGTCAATGCTGTCTGCAAAACTCCT 258

RESULT 45
US-09-925-300-421
; Sequence 421, Application US/09925300
; Patent No. US20020151691A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 421
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-421

Alignment Scores:
Pred. No.: 0.00359 Length: 622
Score: 93.00 Matches: 19
Percent Similarity: 53.19% Conservative: 6
Best Local Similarity: 40.43% Mismatches: 20
Query Match: 28.44% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-925-300-421 (1-622)

Qy 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn---ProProAsn 30
Db 288 AAGGAGGGTTCCTGCCCGGAGCAATTAACCTTCCCGAGCTGGCCTCTGTCTGGGAC 347

; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(264)
US-09-790-264-55

Alignment Scores:
Pred. No.: 0.00161 Length: 471
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservative: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-09-790-264-55 (1-471)

Qy 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProPro----- 29
Db 127 AAACCTGGAGCTGTGCC-----AAGCCTTACCAGAA 159
Qy 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 160 AGTGTGGAAATTGTTGATCAATGCTCAGGAGATGATCCTCCCTGGCAACATGAAG 219
Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 220 TGCTGTAGCAATAGCTGTGGTCAATGCTGTCTGCAAAACTCCT 258

RESULT 46
US-09-924-340-41
; Sequence 41, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 41
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..90
; NAME/KEY: CDS
; LOCATION: 91..462
; NAME/KEY: 3'UTR
; LOCATION: 463..643
; NAME/KEY: polyA signal
; LOCATION: 607..612
; NAME/KEY: polyA site
; LOCATION: 628..643
US-09-924-340-41

Alignment Scores:
Pred. No.: 0.00375 Length: 643
Score: 93.00 Matches: 19
Percent Similarity: 53.19% Conservative: 6
Best Local Similarity: 40.43% Mismatches: 20
Query Match: 28.44% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-924-340-41 (1-643)

Qy 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn---ProProAsn 30
Db 316 AAGGAGGGTTCCTGCCCGGAGCAATTAACCTTCCCGAGCTGGCCTCTGTCTGGGAC 375
Qy 31 ArgCysLeuLysAspThrAspCysProGlyIleLysLysCysGluGlySerCysGly 50
Db 376 CAGTGCCAGGTGGACAGCCAGTGTCTGCGCCAGATGAATGCTGCCGCAATGGCTGTGGG 435
Qy 51 ---MetAlaCysPheValPro 56
Db 436 AAGGTGTCTGTGTCTACTCCC 456

RESULT 47
US-09-992-600A-41
; Sequence 41, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
```



```
Alignment Scores:
Pred. No.: 0.0891 Length: 6314
Score: 92.50 Matches: 18
Percent Similarity: 40.38% Conservative: 3
Best Local Similarity: 34.62% Mismatches: 20
Query Match: 28.29% Indels: 11
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-954-456-319 (1-6314)

QY 12 LysProGlySerCysProfile-----IleLeuIleArgCysAlaMet 25
Db 538 AAGCAGGGGACTGTCGGCTCTGAGAAAGCCAGTGGATTTCGGCGCCCTGTGT--- 594
QY 26 LeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysCysCys 45
Db 595 -----GAAAGCTCGAAGTTGACAATGAGTCTCTGGGGTGAAGAAATGTGT 642
QY 46 GluglySerCysGlyMetAlaCysPheValProGln 57
Db 643 TCGAATGGGTGTGGACACACCTGTCAAGTACCCCAAG 678

RESULT 50
US-09-954-456-1230
; Sequence 1230, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1230
; LENGTH: 6314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1230

Alignment Scores:
Pred. No.: 0.0891 Length: 6314
Score: 92.50 Matches: 18
Percent Similarity: 40.38% Conservative: 3
Best Local Similarity: 34.62% Mismatches: 20
Query Match: 28.29% Indels: 11
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-954-456-1230 (1-6314)

QY 12 LysProGlySerCysProfile-----IleLeuIleArgCysAlaMet 25

Alignment Scores:
Pred. No.: 0.0471 Length: 753
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-852-797-34 (1-753)

QY 5 ValLysGlyProValSerThrIys-----ProGlySerCysProIleIleLeu 21
Db 85 GTCAGGACCTGGTCTGACTGATTTCCAGGAGATGTCCTCCAAATCAGAGAA 144
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGGACAGACAATCCAGGACAAC 201
QY 42 LysIysCysCysGluGlySerCysGlyMetAlaCys 53
```

Db 202 AAGAAGTGTGTCTTCAGCTCGGAGAAAATGT 237
RESULT 52
US-09-853-161-34
; Sequence 34, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-161-34
Alignment Scores:
Pred. No.: 0.0471 Length: 753
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 10 Gaps: 2
09-833799-13B (1-57) x US-09-853-161-34 (1-753)
Qy 5 ValysGlyProValSerThrLys-----ProGlySerCysProIleLeuIle 21
|||:|||||
Db 85 GTCCAGGACCTGCTCTGACTGATTGGTTATTTCCAGGAGATGTCCTCCAAAATCAGAGAA 144
Qy 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
|||:|||||
Db 145 GAATGTGAA---TTCCAGAAAGGAGTGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201
RESULT 53
US-09-852-659A-34
; Sequence 34, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12

; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-34
Alignment Scores:
Pred. No.: 0.0471 Length: 753
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 10 Gaps: 2
09-833799-13B (1-57) x US-09-852-659A-34 (1-753)
Qy 5 ValysGlyProValSerThrLys-----ProGlySerCysProIleLeuIle 21
|||:|||||
Db 85 GTCCAGGACCTGCTCTGACTGATTGGTTATTTCCAGGAGATGTCCTCCAAAATCAGAGAA 144
Qy 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
|||:|||||
Db 145 GAATGTGAA---TTCCAGAAAGGAGTGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201
Qy 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53
|||:|||||
Db 202 AAGAAGTGTGTCTTCAGCTCGGAGAAAATGT 237
RESULT 54
US-09-852-797-50
; Sequence 50, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12

```

; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-797-50

```

```

Alignment Scores:
Pred. No.: 0.0496 Length: 783
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 9 Gaps: 2

```

09-833799-13B (1-57) x US-09-852-797-50 (1-783)

```

QY 5 ValLysGlyProValSerThrLys-----ProGlySerCysProIleIleLeuIle 21
Db 85 GTCCAGGGACCTGGTCTGACTGATTGGTTATTTCCAGGAGATGTCCTCAAAATCAGAGAA 144
QY 22 ArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53
Db 202 AAGAAGTGTGTGTCTTCAGCTCGCGAAAAAATGT 237

```

```

RESULT 55
US-09-853-161-50
; Sequence 50, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19

```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-161-50

```

```

Alignment Scores:
Pred. No.: 0.0496 Length: 783
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 10 Gaps: 2

```

09-833799-13B (1-57) x US-09-853-161-50 (1-783)

```

QY 5 ValLysGlyProValSerThrLys-----ProGlySerCysProIleIleLeuIle 21
Db 85 GTCCAGGGACCTGGTCTGACTGATTGGTTATTTCCAGGAGATGTCCTCAAAATCAGAGAA 144
QY 22 ArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53
Db 202 AAGAAGTGTGTGTCTTCAGCTCGCGAAAAAATGT 237

```

```

RESULT 56
US-09-852-659A-50
; Sequence 50, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19

```

; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-50

Alignment Scores:

Pred. No.: 0.0496 Length: 783
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 25
Best Local Similarity: 40.38% Mismatches: 2
Query Match: 25.99% Indels: 4
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-852-659A-50 (1-783)

QY 5 ValysGlyProValSerThrLys-----ProGlySerCysProIleLeuLeu 21
|||:|||||
Db 85 GTCCAGGACCTGCTGCTGACTGATTGTTATTTCCAGGAGATGTCCTCAAAATCAGAGAA 144
|||:|||||
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
|||:|||||
Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGACAGACAATGCCAGGACAAC 201
|||:|||||
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53
|||:|||||
Db 202 AAGAAGTGTGTCTTCAGCTCGGGAATAAATGT 237
|||:|||||

RESULT 57

US-09-852-659A-121
; Sequence 121, Application US/09852659A
; Patent No. US2002007287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-121

Alignment Scores:

Pred. No.: 0.064 Length: 354
Score: 80.50 Matches: 17
Percent Similarity: 45.00% Conservative: 1
Best Local Similarity: 42.50% Mismatches: 21
Query Match: 24.62% Indels: 1
DB: 10 Gaps: 1

09-833799-13B (1-57) x US-09-852-659A-121 (1-354)

QY 14 GlySerCysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCysLeu 33
|||:|||||
Db 43 GGGAGATGTCCCAAAATCAGAGAAATGTGAA---TTCCAAGAAAGGATGTGTGTAC 99
|||:|||||
QY 34 LysAspThrAspCysProGlyIleLysCysCysGluGlySerCysGlyMetAlaCys 53
|||:|||||
Db 100 AAGGACAGACAATGCCAGGACACACAGAAGTGTGTCTTCAGCTCGGGAATAAATGT 159
|||:|||||

RESULT 58

US-09-833-381-747/c
; Sequence 747, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 747
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(334)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-747

Alignment Scores:

Pred. No.: 0.106 Length: 334
Score: 78.50 Matches: 20
Percent Similarity: 45.16% Conservative: 8
Best Local Similarity: 32.26% Mismatches: 21
Query Match: 24.01% Indels: 13
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-833-381-747 (1-334)

QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleuIle 21
||||:|||||
Db 300 AAGAAGCATTTGGCGGGCCCATTTTCGCCCAAGGAAGGCTCATCGGGGAAGAGACCCCTC 241
||||:|||||
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly--- 40
||||:|||||
Db 240 AGCGTCGCTGCTCCTGCCCTCCCTGAACGCTGT-----CCTTGTCCGACTTGG 193
||||:|||||
QY 41 -----IleLysLysCysCysGluGlySerCysGlyMetAla 52
||||:|||||
Db 192 AGCCCGCGCGGAGCGCTTCATGTTGGGAATAACTGCCCCAGCGCTCGCGGCGCGCG 133
||||:|||||
QY 53 CysPhe 54
Db 132 TCCTTC 127
|||

RESULT 59

US-09-819-136-1
; Sequence 1, Application US/09819136
; Patent No. US20020146789A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.

Downloaded from <http://ajph.org/> on November 10, 2014


```
Db 535 CTTAAAGGCGGCTCCACCATCTTCCCGAGCTTCCTGGTGTCCCATCTTCCTTAGT 476
      :::::
Qy 21 eArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
      :::::
Db 475 CAGACTTGGAATTTTCCAACTGGAAGAGGCTGTGGAAAGCATCAAAACTGTCTCTGC 418
      :::::
RESULT 65
US-09-933-267A-1/c
; Sequence 1, Application US/09933267A
; Patent No. US20020123095A1
; GENERAL INFORMATION:
; APPLICANT: Kalush, Francis et al.
; TITLE OF INVENTION: Estrogen receptor alpha variants and
; TITLE OF INVENTION: methods of detection thereof
; FILE REFERENCE: CLO00258C14
; CURRENT APPLICATION NUMBER: US/09/933,267A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/160626
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 60/183756
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/692414
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/768184
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 09/804076
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/826314
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: human
US-09-933-267A-1
Alignment Scores:
Pred. No.: 2,05e+04 Length: 465237
Score: 69.50 Matches: 21
Percent Similarity: 48.21% Conservative: 6
Best Local Similarity: 37.50% Mismatches: 25
Query Match: 21.25% Indels: 4
DB: 10 Gaps: 1
09-833799-13B (1-57) x US-09-933-267A-1 (1-465237)
Qy 1 AlaGlnGluProValys-----GlyProValSerThrLysProGly-SerCysPr 17
      :::::
Db 261507 GCACGCCGCCGCTCCGGAGGAGGAGTGGGGGGTCCAGCCCCCGCTGGCCAGCTGCC 261448
      :::::
Qy 17 ollelleLeulleAtqCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAs 37
      :::::
Db 261447 AGTCCGGAGGAGTGGGGCGCTCTCCCGAGCGCCCTACTGGGAAGTGAGGAGCCCT 261388
      :::::
Qy 37 pCysProGlyIleLysLysCysCysGlySerCysGlyMetAla 52
      :::::
Db 261387 CTGCCCGCCAGCGCCCGCTCCGGAGGAGGAGTGGGGTTCAGCC 261342
      :::::
RESULT 66
US-09-864-761-18387
; Sequence 18387, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18387
; LENGTH: 148
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050348.19
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EST_HUMAN HIT: BF197609.1, EVALUE 5.00e-78
; OTHER INFORMATION: NT HIT: AB032950.1, EVALUE 1.10e+00
; OTHER INFORMATION: SWISSPROT HIT: O46655, EVALUE 5.00e-05
US-09-864-761-18387
Alignment Scores:
Pred. No.: 0,567 Length: 148
Score: 69.00 Matches: 10
Percent Similarity: 56.00% Conservative: 4
Best Local Similarity: 40.00% Mismatches: 11
Query Match: 21.10% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x US-09-864-761-18387 (1-148)
Qy 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
      :::::
Db 13 TGTGTGATGATGAGATTTCTCAAGTGGAGAAAAATTTTGAAGTCAAGCTGTGGCCGC 72
      :::::
Qy 52 AlaCysPheValPro 56
      :::::
Db 73 TTCTGTGTCCCACCA 87
      :::::
```



```
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225

Alignment Scores:
Pred. No.:      29.4      Length:      2108
Score:          67.50    Matches:      18
Percent Similarity: 46.67% Conservative:   3
Best Local Similarity: 40.00% Mismatches:    17
Query Match:     20.64% Indels:         7
DB:              10      Gaps:           3

09-833799-13B (1-57) x US-09-962-832-225 (1-2108)
QY 10 SerThrLysProGlySerCysProIleLeuLeuArgCysAlaMetLeuAenProPro 29
Db 1042 TCTCCAGCTGTCCAGTTGCCCTTCTGTGTCTCCAGATGC-----TTAGGCTGCCCC 989
QY 30 AsnArgCysLeuLysAspThrAspCysProGlyLeuLysCysCysGludGlySerCys 49
Db 988 TCCTGTCTCTCCAGTGCTTCAGTGGCCCC-----ATCTGTCTGTCTGGGAGCTCT 938
QY 50 GlyMetAlaCysPhe 54
Db 937 GGC-----TGC TTC 929

RESULT 70
US-09-962-832-154
; Sequence 154, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 302250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-154

Alignment Scores:
Pred. No.:      2.39e+04      Length:      302250
Score:          67.00    Matches:      22
Percent Similarity: 43.33% Conservative:   4
Best Local Similarity: 36.67% Mismatches:    19
Query Match:     20.43% Indels:        15
DB:              10      Gaps:           3

09-833799-13B (1-57) x US-09-962-832-154 (1-302250)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuLeuArgCys 23
Db 291823 CCAGTCAAGTTCCCTGTAGTGTA-----AGCTGCCCTTCGCCCTTTTCAAATGT 291873
QY 24 AlaMetLeuAenProProAenArgCysLeu-LysAspThrAspCysProGlyLeuLysLy 43
```

[illegible]

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 6116 TTCAGATGC-----ATCTGCCCCCAGGGTATGAAGTAAAAAGCGAGAACTGCATTGAT 6169
QY 41 IleLysLysCysCysGlu-----GlySerCys----- 49
Db 6170 ATAAATGAATGTGATGAAGATCCCAACATTGTCTTTTGGTTCTGTACTAATACTCCA 6229
QY 50 ---GlyMetAlaCysPheValPro 56
Db 6230 GGGGGCTTCCAGTGCCTCTGCCCC 6253

RESULT 75

US-10-044-090-47
; Sequence 47, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bardman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 47
; LENGTH: 10284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1098479.1
US-10-044-090-47

Alignment Scores:

Pred. No.:	370	Length:	10284
Score:	66.00	Matches:	22
Percent Similarity:	41.18%	Conservative:	6
Best Local Similarity:	32.35%	Mismatches:	24
Query Match:	20.18%	Indels:	16
DB:	12	Gaps:	4

09-833799-13B (1-57) x US-10-044-090-47 (1-10284)

QY 3 GluProValLysGlyProValSerThrLysProGlySerCysProIle-----Leu 20
Db 6056 GAGTGTGTCGCCCTTCCCGGCTTCTGCTCTCCIGGTACCTGTCAGAAATTGGAGGGATCC 6115
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 6116 TTCAGATGC-----ATCTGCCCCCAGGGTATGAAGTAAAAAGCGAGAACTGCATTGAT 6169
QY 41 IleLysLysCysCysGlu-----GlySerCys----- 49
Db 6170 ATAAATGAATGTGATGAAGATCCCAACATTGTCTTTTGGTTCTGTACTAATACTCCA 6229
QY 50 ---GlyMetAlaCysPheValPro 56
Db 6230 GGGGGCTTCCAGTGCCTCTGCCCC 6253

Search completed: February 15, 2003, 22:01:09
Job time : 181 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 21:14:47 ; Search time 704.5 Seconds
(without alignments)
8509.834 Million cell updates/sec

Title: 09-833799-13A

Perfect score: 206

Sequence: 1 AATTCGAGTCGGTACCATA.....GTTTCGTTCCCAATAATAG 206

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.to.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	206	100.0	206	6	A31086 DNA fragmen
2	202	98.1	206	6	A31087 DNA fragmen
3	127	61.7	8598	6	E16036 cDNA encodi
4	125.4	60.9	177	6	196047 Sequence 3
5	125.4	60.9	177	6	196049 Sequence 7
6	125.2	60.8	480	6	AX397178 Sequence
7	124.6	60.5	434	6	AX041087 Sequence
8	124.6	60.5	478	9	HSANTLEUP
9	124.6	60.5	504	6	A31090
10	124.6	60.5	571	6	I74316
11	124.6	60.5	582	9	BC010952 Homo sapi
12	123.8	60.1	177	6	I96048
13	123.2	59.8	1878	9	HUMELAFIN
14	123.2	59.8	2309	6	AX333644
15	123.2	59.8	2309	6	AX334316
16	123.2	59.8	2309	9	HUMPREELAS
17	123.2	59.8	2309	9	S58717 pre-elafin=
18	123.2	59.8	99747	9	HSJ172H20
19	123	59.7	347	6	A31089
20	121	58.7	321	6	I74313
21	120.8	58.6	171	6	A31088
22	84.6	41.1	573	4	BOT223216
23	75.2	36.5	3693	4
24	7.				
25	7				
26	73.1				
27	73.1				
28	71.8				
29	65				
30	68				
31	67.8				
32	67.6				
33	67				
34	66.2				
35	66.2				
36	66.2				
37	66.2				
38	66				
39	64				
40	63				
41	60.4				
42	54.6				
43	43.4				
44	42.6				
45	39				
46	38.4				
47	35.6				
48	35.4				
49	35.2				
50	34.8				
51	34.6				
52	34.6				
53	34.6				
54	34.6				
55	34.2				
56	34				
57	34				
58	33.8				
59	33.8				
60	33.6				
61	33.6				
62	33.6				
63	33.6				
64	33.4				
65	33.4				

default

ALIGNMENTS

FEATURES	SEQUENCE	Location/Qualifiers
source	1..206	
	/organism="synthetic construct"	
	/db xref="taxon:32630"	
BASE COUNT	51 a 46 c 44 g 65 t	
ORIGIN		
Query Match	100.0%;	Score 206; DB 6; Length 206;
Best Local Similarity	100.0%;	Pred. No. 2.5e-53;
Matches 206;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 AATTCGAGCTCGGTACCATCTGCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTA 60	
Db	1 AATTCGAGCTCGGTACCATCTGCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTA 60	
QY	61 CTAAGCCAGGTTCTTGTTCCTATTATCTTGAATTCGTTGCGTATGTTAAACCCACCTAAC 120	

Db	61	CTAAGCAGGTTCTTGTCCTATTATCTTGATTGGCTGATGTTAAACCCACCTAAC	120
Qy	121	GTGTGTTGAAGGACACTGATGTCCAGGATCAAAAGTGCTGCAAGGTTCCCTCGGTA	180
Db	121	GTGTGTTGAAGGACACTGATGTCCAGGATCAAAAGTGCTGCAAGGTTCCCTCGGTA	180
Qy	181	TGGCTGTGTTTCGTTCCACAATAAG 206	
Db	181	TGGCTGTGTTTCGTTCCACAATAAG 206	
RESULT 2			
A31087/c			
LOCUS	A31087	206 bp	DNA
DEFINITION	DNA fragment from patent EP0402068.		
ACCESSION	A31087		
VERSION	A31087.1	GI:1247248	
KEYWORDS	synthetic construct.		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 206)		
AUTHORS	Chrisophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.		
TITLE	Polypeptides and polypeptide analogues with inhibitory activity		
JOURNAL	Patent: EP 0402068-A 28 12-DEC-1990;		
FEATURES	IMPERIAL CHEMICAL INDUSTRIES PLC		
source	Location/Qualifiers		
BASE COUNT	1...206		
ORIGIN	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	64 a	45 c	47 g 50 t
Query Match	98.1%;	Score 202;	DB 6; Length 206;
Best Local Similarity	100.0%;	Pred. No. 4.4e-52;	
Matches 202;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	5	CGAGGTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGGTCCTGTGTTACTAA	64
Db	206	CGAGTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGGTCCTGTGTTACTAA	147
Qy	65	GCAGGTCCTTGCTCTATATCTTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTG	124
Db	146	GCAGGTCCTTGCTCTATATCTTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTG	87
Qy	125	TTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGCTGTAAGGTTCTTCGGTATGCG	184
Db	86	TTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGCTGTAAGGTTCTTCGGTATGCG	27
Qy	185	TGTTTCGTTCCACAATAAG 206	
Db	26	TGTTTCGTTCCACAATAAG 5	
RESULT 3			
E16036			
LOCUS	E16036	8598 bp	DNA
DEFINITION	cDNA encoding improved Pichia elafin.		
ACCESSION	E16036		
VERSION	E16036.1	GI:5710719	
KEYWORDS	JP 1998127292-A/1.		
SOURCE	Pichia pastoris.		
ORGANISM	Pichia pastoris		
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
AUTHORS	Taniyama,M., Yamamoto,T., Okawa,N. and Zushi,M.		
TITLE	ELAFIN-EXPRESSION VECTOR AND PRODUCTION OF ELAFINS BY UTILIZING		
JOURNAL	THE SAME		
Patent:	JP 1998127292-A 1 19-MAY-1998;		
TSUMURA & CO			
COMMENT	OS Pichia pastoris		

PN JP 1998127292-A/1
PD 19-MAY-1998
PF 31-OCT-1996 JP 1996304233
PI TANIYAMA MITSUE, YAMAMOTO TAKASHI, OKAWA NORIYUKI, PI ZUSHI
MAKOTO
PC C12N15/09, C07H21/04, C07K14/39, C12N1/19, C12P21/02, C12N1/19, PC
C12R1/84)
PC (C12P21/02, C12R1/84);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1. .8598
FT /organism="Pichia pastoris"
FT /clone="pPIC9/ELP25L"
FT CDS 949. .1377
FT /product="improved pichia elafin" FT
FT sig_peptide 949. .1203
FT mat_peptide 1204. .1374
FT /product="improved pichia elafin" FT
FT misc_feature replace(1276, .1278, 'atg')
FT /note="native Pichia elafin".
FEATURES
source Location/Qualifiers
1. .8598
/organism="Pichia pastoris"
/db_xref="taxon:4922"
BASE COUNT 2306 a 2025 c 1920 g 2347 t
ORIGIN
Query Match 61.7%; Score 127; DB 6; Length 8598;
Best Local Similarity 82.9%; Pred. No. 9.1e-29;
Matches 145; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 30 GCTCAAGAACCGATTAAAGTCTGTCTACTAAAGCCAGGTTCTTGTCTTATTATCTTG 89
DB 1204 GCTCAAGAACCGATTAAAGTCTGTCTACTAAAGCCAGGTTCTTGTCTTATTATCTTG 1263
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGGACACTGATTGTCAGGT 149
DB 1264 ATTCGTTGCGCTATGTTGAACCCGCGGACCGTTGTTGTTGAAGGACACTGATTGTCAGGT 1323
QY 150 ATCAAAAGTCTGTGAAGTCTTCTCGGATGCGTGTGTTTCTGTTCCACATAAT 204
DB 1324 ATCAAAAGTCTGTGAAGTCTTCTCGGATGCGTGTGTTTCTGTTCCACATAAT 1378
RESULT 4
LOCUS 196047
DEFINITION Sequence 3 from patent US 5734014.
ACCESSION 196047
VERSION 196047.1 GI:3940517
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima, Y., Okawa, N., Yoshida, M., Amagaya, S. and Kajii, A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 3 31-MAR-1998;
FEATURES
source Location/Qualifiers
1. .177
/organism="unknown"
BASE COUNT 39 a 49 c 48 g 41 t
ORIGIN
Query Match 60.9%; Score 125.4; DB 6; Length 177;
Best Local Similarity 82.3%; Pred. No. 3.1e-28;
Matches 144; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 30 GCTCAAGAACCGATTAAAGTCTGTCTACTAAAGCCAGGTTCTTGTCTTATTATCTTG 89

DB 1 GCACAGGAACCGATTAAAGTCCGGTGTGCGACCAACCGGCTCTTGTCCCGATTATCTTG 60
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGGACACTGATTGTCAGGT 149
DB 61 ATCCGCTCGGCTTGTCTGAACCCGCGGACCGTTGTCTGAAAGACACTGACTGCCCGGT 120
QY 150 ATCAAAAGTCTGTGAAGTCTTCTCGGATGCGTGTGTTTCTGTTCCACATAAT 204
DB 121 ATCAAAAGTCTGTGAAGTCTTCTCGGATGCGTGTGTTTCTGTTCCACATAAT 175
RESULT 5
LOCUS 196049
DEFINITION Sequence 7 from patent US 5734014.
ACCESSION 196049
VERSION 196049.1 GI:3940519
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima, Y., Okawa, N., Yoshida, M., Amagaya, S. and Kajii, A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 7 31-MAR-1998;
FEATURES
source Location/Qualifiers
1. .177
/organism="unknown"
BASE COUNT 40 a 49 c 47 g 41 t
ORIGIN
Query Match 60.9%; Score 125.4; DB 6; Length 177;
Best Local Similarity 82.3%; Pred. No. 3.1e-28;
Matches 144; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 30 GCTCAAGAACCGATTAAAGTCTGTCTACTAAAGCCAGGTTCTTGTCTTATTATCTTG 89
DB 1 GCACAGGAACCGATTAAAGTCCGGTGTGCGACCAACCGGCTCTTGTCCCGATTATCTTG 60
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGGACACTGATTGTCAGGT 149
DB 61 ATCCGCTCGGCTATGTTGAACCCGCGGACCGTTGTCTGAAAGACACTGACTGCCCGGT 120
QY 150 ATCAAAAGTCTGTGAAGTCTTCTCGGATGCGTGTGTTTCTGTTCCACATAAT 204
DB 121 ATCAAAAGTCTGTGAAGTCTTCTCGGATGCGTGTGTTTCTGTTCCACATAAT 175
RESULT 6
LOCUS AX397178/c
DEFINITION Sequence 1393 from Patent WO212328.
ACCESSION AX397178
VERSION AX397178.1 GI:21067925
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Fukuyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL King, G.E., Meagher, M.J., Xu, J., and Secrist, H.
FEATURES
source Compositions and methods for the therapy and diagnosis of colon
cancer
Patent: WO 0212328-A 1393 14-FEB-2002;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .480
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 113 a .105 c 139 g 120 t
ORIGIN

```

Query Match      60.8%; Score 125.2; DB 6; Length 480;
Best Local Similarity 81.0%; Pred. No. 3.5e-28;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTTAAAGCTCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 371 AAGCGCAAGAGCCAGTCAAGAGTCCAGTNTCCACTAAGCCCTGGCTCCTGCCCATTAATCT 312
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 88 TGATTCGTTGCGCTATGTTAAACCCACTTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 311 TGATTCGTTGCGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACACTGATGCCAG 252
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 251 GAATCAAGAAGTGTGTAAGGCTTTGCGGATGGCTGTTTCTGCCAGTGAGAG 193
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
AX041087/c      434 bp      DNA      linear      PAT 23-NOV-2000
LOCUS           AX041087
DEFINITION      Sequence 5 from Patent WO0065053.
ACCESSION       AX041087
VERSION         AX041087.1 GI:11340657
KEYWORDS        human.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 434)
AUTHORS         Wang, T. and Dillon, D.C.
TITLE           Compositions and methods for therapy and diagnosis of head/neck and
lung squamous cell carcinoma
JOURNAL         Patent: WO 0065053-A 5 02-NOV-2000;
CORIXA CORPORATION (US)
FEATURES        Location/Qualifiers
source          1..434
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      103 a 103 c 133 g 94 t 1 others
ORIGIN

Query Match      60.5%; Score 124.6; DB 6; Length 434;
Best Local Similarity 81.0%; Pred. No. 5.4e-28;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 320 AAGCGCAAGAGCCAGTCAAGTCCAGTCTCCACTAAGCCCTGGCTCCTGCCCATTAATCT 261
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 88 TGATTCGTTGCGCTATGTTAAACCCACTTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 260 TGATTCGTTGCGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACACTGACTGCCAG 201
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 GAATCAAGAAGTGTGTAAGGCTTTGCGGATGGCTGTTTCTGCCAGTGAGAG 142
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
HSANTLEUP      478 bp      mRNA      linear      PRI 08-JAN-1998
LOCUS           HSANTLEUP
DEFINITION      H.sapiens encoding skin-derived antileukoproteinase.
ACCESSION       Z18538
VERSION         Z18538.1 GI:28711
KEYWORDS        antileukoproteinase; elafin; elastase inhibitor; proteinase
inhibitor; psoriasis; SKALP; transglutaminase substrate.
SOURCE          Homo sapiens.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 478)
AUTHORS         Molhuizen, H.O., Alkemade, H.A., Zeeuwen, P.L., de Jongh, G.J.,
Wieringa, B. and Schalkwijk, J.

```

```

TITLE           SKALP/elafin: an elastase inhibitor from cultured human
keratinocytes. Purification, cDNA sequence, and evidence for
transglutaminase cross-linking
J. Biol. Chem. 268 (16), 12028-12032 (1993)
93280175
7685029
MEDLINE
PUBMED
REFERENCE       2 (bases 1 to 478)
AUTHORS         Molhuizen, H.O.F.
TITLE           Direct Submission
JOURNAL         Submitted (20-NOV-1992) Molhuizen H.O.F., Academic Hospital,
Dermatology, Nijmegen, The Netherlands, 6500 HB
FEATURES        Location/Qualifiers
source          1..478
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pGESKA"
/cell_type="keratinocyte"
/clone_lib="HD110b Clonetechn"
1..478
/partial
/citation=[1]
3..356
/function="proteinase inhibitor"
/codon_start=1
/evidence-experimental
/product="skin-derived antileukoproteinase"
/protein_id="CAA79223.1"
/db_xref="GI:28712"
/db_xref="SWISS-PROT:P19957"
/translation="MRASSFLIVVFLIAGTLVLEAAVTGVPYKQDVTYKGRVFPNGQ
DPVKGQSVKQDKVKAQEPVKGPVSTKPGSPILLIRCAMLNPPNRLKLDTCPCGIR
KCCGSGCMACFVPQ"
3..68
sig_peptide    69..353
mat_peptide    69..353
/partial
/product="skin-derived antileukoproteinase"
/function="proteinase inhibitor"
BASE COUNT      102 a 142 c 119 g 115 t
ORIGIN

Query Match      60.5%; Score 124.6; DB 9; Length 478;
Best Local Similarity 81.0%; Pred. No. 5.4e-28;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AAGCGCAAGAGCCAGTCAAGGTTCCAGTCTCCACTAAGCTGGCTCCTGCCCATTAATCT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 88 TGATTCGTTGCGCTATGTTAAACCCACTTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TGATTCGTTGCGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACACTGACTGCCAG 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 GAATCAAGAAGTGTGTAAGGCTTTGCGGATGGCTGTTTCTGCCAGTGAGAG 359
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
A31090          504 bp      DNA      linear      PAT 21-AUG-1995
LOCUS           A31090
DEFINITION      DNA fragment from patent EP0402068.
ACCESSION       A31090
VERSION         A31090.1 GI:1249303
KEYWORDS        synthetic construct.
SOURCE          synthetic construct.
ORGANISM        artificial sequences.
REFERENCE       1 (bases 1 to 504)
AUTHORS         Christophers, E., Schroeder, J.M., Pioli, D., Wiedow, O. and Edge, M.D.
TITLE           Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL         Patent: EP 0402068-A 31 12-DEC-1990;

```

IMPERIAL CHEMICAL INDUSTRIES PLC
 Location/Qualifiers
 1..504
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 110 a 120 g 122 t
 ORIGIN

Query Match 60.5%; Score 124.6; DB 6; Length 504;
 Best Local Similarity 81.0%; Pred. No. 5.4e-28;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGTTCCTGTCCTATTATCT 87
 DB 156 AAGCGCAAGAGCCAGTCAAGGTCCTCAAGCTGCTCCCTGCTCCGCCATTATCT 215

QY 88 TGATTCCTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147
 DB 216 TGATCCGGTGGCCATGTTGAATCCCTCAACCGCTGTTGAAGATATCTGACTGCCAG 275

QY 148 GTATCAAAAAGTCTGTGAAGGTCCTGCGGTATGCTTGTTCGTTCCACAATAATAG 206
 DB 276 GAATCAAGAAGTCTGTGAAGGCTTTCGGGATGCGCTTTCGTTCCCGAGTGAGAG 334

RESULT 10
 I74316
 LOCUS I74316 571 bp DNA linear PAT 03-APR-1998
 DEFINITION Sequence 14 from patent US 5688641.
 ACCESSION I74316
 VERSION I74316.1 GI:3010457
 KEYWORDS Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 571)
 AUTHORS Sager R., Zou Z., Lee S., Whan, and Tomasetto, C. Laure.
 TITLE Cancer diagnosis using nucleic acid hybridization
 JOURNAL Patent: US 5688641-A 14 18-NOV-1997;
 FEATURES Location/Qualifiers
 1..571
 /organism="unknown"
 BASE COUNT 138 a 167 c 130 g 136 t
 ORIGIN

Query Match 60.5%; Score 124.6; DB 6; Length 571;
 Best Local Similarity 81.0%; Pred. No. 5.3e-28;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGTTCCTGTCCTATTATCT 87
 DB 197 AAGCGCAAGAGCCAGTCAAGGTCCTGCTCCACTAAGGCTGCTCCCTGCCCATTTATCT 256

QY 88 TGATTCCTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147
 DB 257 TGATCCGGTGGCCATGTTGAATCTCTCCCTCAACCGTCTTGAAGATATCTGACTGCCAG 316

QY 148 GTATCAAAAAGTCTGTGAAGGTCCTGCGGTATGCTTGTTCGTTCCACAATAATAG 206
 DB 317 GAATCAAGAAGTCTGTGAAGGCTTTCGGGGATGSCCTGTTTCGTTCCCGAGTGAGAG 375

RESULT 11
 BC010952
 LOCUS BC010952 582 bp mRNA linear PRI 25-JUL-2001
 DEFINITION Homo sapiens, Similar to protease inhibitor 3, skin-derived (SKALP), clone MGC:13613 IMAGE:4083155, mRNA, complete cds.
 ACCESSION BC010952
 VERSION BC010952.1 GI:15012094
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT

FEATURES
 source

CDS

BASE COUNT
 ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 582;
 Best Local Similarity 81.0%; Pred. No. 5.3e-28;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGTTCCTGTCCTATTATCT 87
 DB 204 AAGCGCAAGAGCCAGTCAAGGTCCTGCTCCACTAAGGCTGCTCCCTGCCCATTTATCT 263

QY 88 TGATTCCTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147
 DB 264 TGATCCGGTGGCCATGTTGAATCTCTCCCTCAACCGTCTTGAAGATATCTGACTGCCAG 323

QY 148 GTATCAAAAAGTCTGTGAAGGTCCTGCGGTATGCTTGTTCGTTCCACAATAATAG 206
 DB 324 GAATCAAGAAGTCTGTGAAGGCTTTCGGGGATGSCCTGTTTCGTTCCCGAGTGAGAG 382

RESULT 12

I96048
 LOCUS I96048 177 bp DNA linear PAT 01-DEC-1998
 DEFINITION Sequence 5 from patent US 5734014.
 ACCESSION I96048
 VERSION I96048.1 GI:3940518
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 582)
 Strausberg, R.
 Direct Submission
 Submitted (23-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-re@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 19 Row: d Column: 7
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4505786.

Location/Qualifiers

1..582
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:13613 IMAGE:4083155"
 /tissue_type="Brain, glioblastoma"
 /clone_lib="NIH MGC_57"
 /lab_host="DH10B"
 /note="Vector: pDNR-Lib"
 26..379
 /codon_start=1
 /product="Similar to protease inhibitor 3, skin-derived (SKALP)"
 /protein_id="AAH10952.1"
 /db_xref="GI:15012095"
 /translation="MKASSFLIVVFLIAGTLVLEAAVTGVPVKQDVTVKGRVPFNGQ
 DPKGVQSVKQDKVKAQEPVKPVPKSGCPILLIRCAMLNPPNRCUKLDTDCRGIK
 KCCEGSCGMACEFVPQ"

BASE COUNT 148 a 166 c 134 g 134 t

```

REFERENCE 1 (bases 1 to 177)
AUTHORS Iehima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaji,A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 5 31-MAR-1998;
FEATURES Location/Qualifiers
source
1..177
/organism="unknown"
BASE COUNT 39 a 50 c 48 g 40 t
ORIGIN

Query Match 60.1%; Score 123.8; DB 6; Length 177;
Best Local Similarity 81.7%; Pred. No. 9.7e-28;
Matches 143; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 30 GCTCAAGAACAGTTAAAGGTCCTGCTACTAAGCCAGGCTTCTGTGCTTATTTATCTTG 89
Db 1 GCACAGAACAGTTAAAGGTCCTGCTGCTGACCAACCGGCTCTTGCCCGATTATCTG 60

QY 90 ATTCGCTGGCTATGTTAAACCCACTAACCGTTGTTTGAAGGACACTGATTGCCAGGT 149
Db 61 ATCCGCTGGCTGTCTGAAACCCGCGGAACCGTTGTTCTGAAAGACACTGACTGCCGGGT 120

QY 150 ATCAAAAAGTGTGAAGGTTCTGCGGTATGCTGTTGTTTCCACATATAT 204
Db 121 ATCAAAAAGTGTGCAAGGTTCTTGGCGTATGCGATGCTTCTGTTCCGCGAGT 175

RESULT 13
LOCUS HUMELAFIN 1878 bp DNA linear PRI 14-APR-2000
DEFINITION Human gene for elafin, complete cds.
ACCESSION D13156
VERSION D13156.1 GI:219614
KEYWORDS elafin; elastase inhibitor.
SOURCE Homo sapiens (library: lambda EMBL) DNA, clone lambda hi-G1..
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1878)
AUTHORS Saeki,T., Ito,F., Hagiwara,H., Saito,Y., Kuroki,J., Tachibana,S.
and Hirose,S.
TITLE Primary structure of the human elafin precursor preproelafin
deduced from the nucleotide sequence of its gene and the presence
of unique repetitive sequences in the prosegment
JOURNAL Biochem. Biophys. Res. Commun. 185 (1), 240-245 (1992)
MEDLINE 92287100
REFERENCE 2 (bases 1 to 1878)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1992) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; Ookayama, Meguro-ku,
Tokyo 152, Japan (Tel:03-3726-1111(ex.2226), Fax:03-3729-0335)
COMMENT Submitted (01-SEP-1992) to DDBJ by:
Shigehisa Hirose
Dept. of Biological Sciences
Tokyo Institute of Technology
Ookayama, Meguroku
Tokyo 152
Japan
Phone: 03-3726-1111
Fax: 03-3729-0335.
FEATURES source
1..1878
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="lambda EMBL"
CAAT_signal 127..130
TATA_signal 153..158
exon <181..325
number=1
CDs join(247..325,1185..1459)
/codon_start=1
/product="preproelafin"

/protein_id="BAA02441.1"
/db_xref="GI:219615"
/translation="MRASFLIVVFLIAGTLVLEAAVTGVPVKGDTVKRVPFNQ
DPVKGQSVKQDKVKAQEPVKGPSTKPGSCPIILIRCAMLNPPNRLCLDTPCGIK
KCCGSGCMACFPVQ"
250..324
sig_peptide 1286..1456
mat_peptide /product="elafin"
1185..1460
exon /number=2
1695..1852
exon /number=3
polyA_signal 1847..1852
BASE COUNT 473 a 464 c 491 g 450 t
ORIGIN

Query Match 59.8%; Score 123.2; DB 9; Length 1878;
Best Local Similarity 81.2%; Pred. No. 1.4e-27;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGTCTGTTGCTTAAAGCCAGGTTCTTGTCTTATTTATCT 87
Db 1284 AAGCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAAGCTGGCTCCTGCCCATTTATCT 1343

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 1344 TGATCCGGTGGCGCATGTTGAAATCCCGCTAACCGTGTCTTGAAGATAGTACTGACTGCCAG 1403

QY 148 GTATCAAAAAGTGTGTTGAAGGTTCTGCGGTATGCTGTTGTTTCCACATAA 203
Db 1404 GAATCAAGAGTGTGTTGAAGGCTCTTGGCGATGCGCTGTTTCCGCCAGTGA 1459

RESULT 14
LOCUS AX333644 2309 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 4153 from Patent WO0194629.
ACCESSION AX333644
VERSION AX333644.1 GI:18124363
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4153 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source
1..2309
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 595 a 537 c 603 g 572 t 2 others
ORIGIN

Query Match 59.8%; Score 123.2; DB 6; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.4e-27;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGTCTGTTGCTTAAAGCCAGGTTCTTGTCTTATTTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAAGCTGGCTCTGCCCATTTATCT 1611

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 1612 TGATCCGGTGGCGCATGTTGAATCCCGCTAACCGTGTCTTGAAGATAGTACTGACTGCCAG 1671

QY 148 GTATCAAAAAGTGTGTTGAAGGTTCTGCGGTATGCTGTTGTTTCCACATAA 203
Db 1672 GAATCAAGAGTGTGTTGAAGGCTCTTGGCGGATGGCTGTTTCTGTTCCCGCATGA 1727

```

```

RESULT 15
AX334316
LOCUS
DEFINITION Sequence 4825 from Patent WO0194629.
ACCESSION AX334316
VERSION AX334316.1 GI:18125035
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrihan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4825 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
1..2309
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 595 a 537 c 603 g 572 t 2 others
ORIGIN
Query Match 59.8%; Score 123.2; DB 6; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.4e-27;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Qy 28 ATGCTCAAGAACCAAGTCTGCTGCTCTACTAAGCCAGGTTCTGTCCTATTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCCTGCTCTGCCCATTTATCT 1611
Qy 88 TGATTCCTGCGCATCTTAAACCCACTAACCGTTGTTTGAAGGACACTGATGTCAC 147
Db 1612 TGATCCGGTGGCCATGTTGAATCCCTTAACCGCTGTTGAAGAGATACACTGCTGCC 1671
Qy 148 GTATCAAAAAGTGTGTGAAGGTTCTGCGGTATGCTGTTCTGTTCCACAATAA 203
Db 1672 GAATCAAGAGTGTGTGAAGGTTCTGCGGTATGCTGTTCTGTTCCCAAGTGA 1727

RESULT 16
HUMPREELAS
LOCUS
DEFINITION Homo sapiens elafin precursor, gene, complete cds.
ACCESSION LI0343
VERSION LI0343.1 GI:190337
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Sallenne,J.M. and Silva,A.
TITLE Characterization and gene sequence of the precursor of elafin, an
elastase-specific inhibitor in bronchial secretions
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (4), 439-445 (1993)
MEDLINE
PUBMED 9326929
REMARK 8476637
FEATURES
Location/Qualifiers
1..2309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
/tissue_lib="EMBL-3 (from Clontech)"
1..595
/number=1
398..401
Join(517..595,1453..1727)
/function="elastase-specific proteinase inhibitor"
exon
CAAT_signal
CDS

```

```

/note="elafin has been sequenced at the protein level;
pre-elafin has not; its existence is assumed from its
molecular weight (PAGE analysis); putative"
/codon_start=1
/product="elafin precursor"
/protein_id="AAA36483.1"
/db_xref="GI:190338"
/translation="MRASSFLIVVFLIAGTLVLEAAVTGVPVKQDTPVKRVPFNGQ
DPVKGQVSVKQDKVKAQEPKVPSTKPGSCPILIRCAMLPPNRCCLKDTCPGIK
KCEGSCGMACEVPQ"
1554..1724
/product="elafin"
596..1452
/number=1
1453..1728
/number=2
1729..1961
/number=2
1962..2119
/number=3
2114..2119
polyA_signal
595 a 537 c 603 g 572 t 2 others
ORIGIN
Query Match 59.8%; Score 123.2; DB 9; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.4e-27;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Qy 28 ATGCTCAAGAACCAAGTCTGCTGCTCTACTAAGCCAGGTTCTGTCCTATTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCCTGCTCTGCCCATTTATCT 1611
Qy 88 TGATTCCTGCGCATCTTAAACCCACTAACCGTTGTTTGAAGGACACTGATGTCAC 147
Db 1612 TGATCCGGTGGCCATGTTGAATCCCTTAACCGCTGTTGAAGAGATACACTGCTGCC 1671
Qy 148 GTATCAAAAAGTGTGTGAAGGTTCTGCGGTATGCTGTTCTGTTCCACAATAA 203
Db 1672 GAATCAAGAGTGTGTGAAGGTTCTGCGGTATGCTGTTCTGTTCCCAAGTGA 1727

RESULT 17
S58717
LOCUS
DEFINITION pre-elafin=elastase-specific inhibitor [human, placental, Genomic,
2309 nt].
ACCESSION S58717
VERSION S58717.1 GI:299840
KEYWORDS
SOURCE Homo sapiens placental.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Sallenne,J.M. and Silva,A.
TITLE Characterization and gene sequence of the precursor of elafin, an
elastase-specific inhibitor in bronchial secretions
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (4), 439-445 (1993)
MEDLINE
PUBMED 9326929
REMARK 8476637
FEATURES
Location/Qualifiers
1..2309
/organism="Homo sapiens"
/db_xref="taxon:9606"
Join(398..595,1453..1727,1962..2119)
517..1727
/gene="pre-elafin"
Join(517..595,1453..1727)
/gene="pre-elafin"
/note="elastase-specific inhibitor; This sequence comes

```


repeat_region	/note="HERVL repeat: matches 4. .2665 of consensus" 17140. .17589
repeat_region	/notes="MLT2B repeat: matches 1. .444 of consensus" 17590. .17899
repeat_region	/notes="L1 repeat: matches 4299. .4610 of consensus" 17898. .18682
repeat_region	/notes="L1M1 repeat: matches -1390. -.578 of consensus" 19077. .19213
repeat_region	/notes="L1ME repeat: matches 5694. .5824 of consensus" 20761. .20831
repeat_region	/notes="L1M4b repeat: matches -259. -.184 of consensus" 20853. .21014
repeat_region	/notes="L1MA9 repeat: matches 6102. .6270 of consensus" 21044. .21573
repeat_region	/notes="L1M4b repeat: matches -263. .330 of consensus" 21634. .21829
repeat_region	/notes="L1M4b repeat: matches 330. .524 of consensus" 22891. .23139
repeat_region	/notes="AluX repeat: matches 1. .309 of consensus" 25460. .25629
repeat_region	/notes="L1R16A repeat: matches 201. .378 of consensus" 27612. .27639
repeat_region	/note="14 copies 2 mer ac 100% conserved" 27655. .27986
repeat_region	/notes="L1M3A repeat: matches 202. .531 of consensus" 28959. .29336
repeat_region	/notes="L1M1 repeat: matches 469. .859 of consensus" 29337. .29711
repeat_region	/note="THE1C repeat: matches 1. .371 of consensus" 29712. .31317
repeat_region	/note="THE1C-internal repeat: matches 1. .1580 of consensus" 31318. .31729
repeat_region	/notes="THE1C repeat: matches 1. .371 of consensus" 31730. .31886
repeat_region	/note="L1M1 repeat: matches 859. .987 of consensus" 33209. .34261
repeat_region	/notes="L1M3d repeat: matches 859. .1833 of consensus" 34232. .34942
repeat_region	/notes="L1M1 repeat: matches 1560. .2657 of consensus" 34941. .36117
repeat_region	/note="L1 repeat: matches 3935. .5142 of consensus" 36118. .36412
repeat_region	/notes="AluJo repeat: matches 9. .292 of consensus" 36430. .37302
repeat_region	/notes="L1 repeat: matches 2905. .3780 of consensus" 37299. .37435
repeat_region	/note="L1 repeat: matches 4912. .5038 of consensus" 37434. .37469
repeat_region	/notes="L1P repeat: matches 2850. .2885 of consensus" 38377. .39183
repeat_region	/note="L1M4c repeat: matches 1312. .1666 of consensus" 39261. .42088
repeat_region	/notes="L1PA15 repeat: matches 671. .3485 of consensus" 42057. .43098
repeat_region	/notes="L1PA5 repeat: matches 5104. .6143 of consensus" 43119. .45775
repeat_region	/note="L1PA15 repeat: matches 3467. .6157 of consensus" 45840. .45956
repeat_region	/note="L1ME repeat: matches 5530. .5647 of consensus" 46081. .46114
repeat_region	/note="17 copies 2 mer tc 82% conserved" 46492. .46761
repeat_region	/note="L1ME repeat: matches 5521. .5793 of consensus" 46828. .46915
repeat_region	/note="L1ME3A repeat: matches 5300. .5388 of consensus" 46916. .47210
repeat_region	/note="AluSq repeat: matches 1. .305 of consensus" 47211. .47769
misc_feature	/note="L1ME3A repeat: matches 5388. .5937 of consensus complement (48203. .48496)
repeat_region	/note="match: GSS: Em:AQ275745" 48638. .48679

```

/note="21 copies 2 mer ac 97% conserved"
49535..51104
/genes="PI3"
join(49535..49615,50475..50750,50984..51104)
/genes="PI3"
/product="dJ172H20.3 (protease inhibitor 3, skin-derived
(SKALP))"
/note="match: cDNAs: Em:Z18538 Em:AJ223218 Em:AJ223216
match: ESTs: Em:AA527557 Em:AA583567 Em:AA582866"
/evidence=not experimental
join(49537..49615,50475..50749)

CDS
Query Match      59.8%; Score 123.2; DB 9; Length 93747;
Best Local Similarity 81.2%; Pred. No. 1.3e-27;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGACCAAGTTAAAGGTCCTGTGCTACTACTAAGCCAGGTTCTTGTCCTATTATCT 87
Db 50574 AAGCGCAAGACCCAGTCAAAGGTCAGGTCCTCACTAAGCTGGCTCTGCCCATTTATCT 50633

Qy 88 TGATTCGTTGGCGCTATGTTAAACCCACCTAACCCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 50634 TGATCCGGTGGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAGAGACTGACTGCCAG 50693

Qy 148 GTATCAAAAAGTGCTGAAGGTTCTGCGGTATGCGTGTGTTTCGTTCCACAATAA 203
Db 50694 GAATCAAGAAGTGCTGTGAAGGCTCTTGCGGATGGCGCTGTTTCGTTCCCAAGTGA 50749

RESULT 19
A31089
LOCUS       A31089             347 bp     DNA             linear     PAT 21-AUG-1995
DEFINITION DNA fragment from patent EP0402068.
ACCESSION   A31089
VERSION     A31089.1  GI:1249302
KEYWORDS    .
SOURCE      synthetic construct.
            synthetic construct
            artificial sequences.
            1 (bases 1 to 347)
REFERENCE   1  (bases 1 to 347)
            Christoffers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
            Polypeptides and polypeptide analogues with inhibitory activity
            against human elastase
            Patent: EP 0402068-A 30 12-DEC-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES    Location/Qualifiers
             source                1..347
             organism="synthetic construct"
             db_xref="taxon:32630"
BASE COUNT  65 a 118 c 78 g 84 t                2 others
ORIGIN
Query Match      59.7%; Score 123; DB 6; Length 347;
Best Local Similarity 80.2%; Pred. No. 1.7e-27;
Matches 142; Conservative 2; Mismatches 33; Indels 0; Gaps 0;

Qy 30 GCTCAAGAACCAAGTTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGTCCTATTATCTTG 89
Db 1  GCGCAAGACCAAGTCAAGGTCAGTCTCCACTAAGCTGGCTCTGCGCCCATTTATCTTG 60

Qy 90 ATTGTTGCGGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGT 149
Db 61 ATCCGGTCGCCATGTTGAATCCCCCTAACCGTGTGTTGAAGACTACTGACTGCCAGGA 120

Qy 150 ATCAAAAAGTGCTGAAGGTTCTGCGGTATGCGTGTGTTTCGTTCCACAATAATAG 206
Db 121 ATGAAGAARTCTGTGAAGGCTCTTGCGGGATGGCGTGTTCGTTCCCAAGTAGAG 177

RESULT 20
I74313
LOCUS       I74313             321 bp     DNA             linear     PAT 03-APR-1998
DEFINITION   Sequence 11 from patent US 5688641.

```



```

QY 33 CAAGAACCAAGTTAAAGTCTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGATT 92
|||||
Db 235 CAAGGCCGAATCGGAGTCCACTCTCTCACTAAGCCTGGGTCTCTGCCAGGGTTCTGATC 294
|||||
QY 93 CGTTGGCTATGTTAAACCCACCTAACCTGTTGTTGAAGACACTGATGTCTCAGGTATC 152
|||||
Db 295 CGGTGTGCCATGATGAACCCCTTAACCCGATGCTGAGGATGCTCAGTGCCCGGGGTC 354
|||||
QY 153 AAAAAGTGTGTGAAGGTTCTCTCGGTATGGCTGTTGTTCCCAATAA 203
|||||
Db 355 AAGAAGTGTGTGAAGCTCTTGTGGAGACCTGATGATGATCCCAAGTGA 405
|||||

RESULT 23
LOCUS PIGWAPA 3693 bp DNA linear MAM 14-APR-2000
DEFINITION Pig DNA for elafin, complete cds.
ACCESSION D50319
VERSION D50319.1 GI:1228057
KEYWORDS elafin.
SOURCE Sus scrofa DNA, clone.lib:EMBL SP6/T7 clone:lambda WAP-1.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
Tachibana,S. and Hirose,S.
Accelerated evolution in inhibitor domains of porcine elafin family
members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
96215132
REFERENCE 2 (bases 1 to 3693)
AUTHORS Hirose,S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3693)
AUTHORS Hirose,S.
DIRECT SUBMISSION
TITLE Direct Submission
SUBMITTED (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.citech.ac.jp, Tel:045-924-5726,
Fax:045-924-5824)
FEATURES
source
Location/Qualifiers
1..3693
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="lambda WAP-1"
/clone_lib="EMBL SP6/T7"
<1..710
/number=1
<1..628
join(629..710,1541..1962)
/codon_start=1
/product="elafin"
/protein_id="BAA08854.1"
/db_xref="GI:1228058"
/translation="MRSRSFLVTVVFLICGTLVAAGRIIRPKGKTKKILALVKG
QSPVRGKDVKGQGVKQDGLGSKDPAQLPKDQDLGKGSVKGQDPFKALPD
KLQDPVKAQPAIKRLILITKPGSPRIILICLMVNPNNRCLSDAQCPGLKKCCGFCG
KACMPK"
629..691
1777..1959
/product="elafin"
/number=1
1541..1963
/number=2
join(1963,2445..2601)
1964..2444
/number=2
2141..2402
/note="SINE (PRE-1)"

sig_peptide
mat_peptide
intron
exon
3'UTR
intron
repeat_unit

```

```

exon
2445..2601
/number=3
polyA_signal 2580..2585
polyA_site 2601
repeat_region 3606..3669
/note="TG repeat"
BASE COUNT 798 a 953 c 1112 g 830 t
ORIGIN
Query Match 36.5%; Score 75.2; DB 4; Length 3693;
Best Local Similarity 70.1%; Pred. No. 1.2e-12;
Matches 101; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 60 ACTAAGCCAGGTTCTTGTCTCTATTATCTTGAATTCGTTGCGCTATGTTAAACCCACCTAAC 119
|||||
Db 1819 ACCAAGCCTGGCTCTCTGCCAGGATTCGATCGTTGCTTGATGGTCAATCCCCCTAAC 1878
|||||
QY 120 CGTTGTTTGAAGNACACTGATGTCAGGTATCAAAAGTCTGCTGAAGGTTCTTCGCGGT 179
|||||
Db 1879 AGGTGTTTGAATGATGCTCAGTCCCGAGGCTCAAGAAAGTCTGTGAAGGCTTTTTCGGG 1938
|||||
QY 180 ATGGCTTGTGTTGTTCCACAATAA 203
|||||
Db 1939 AAGCCCTGATGGATCCCAAGTGA 1962
|||||

RESULT 24
LOCUS A31082 74 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide E115 from patent EP0402068.
ACCESSION A31082
VERSION A31082.1 GI:1249296
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 74)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 22 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
Location/Qualifiers
1..74
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 19 a 14 c 18 g 23 t
ORIGIN
Query Match 35.9%; Score 74; DB 6; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.1e-12;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 133 ACATGATGTCAGGTATCAAAAAGTCTGTGAAGTTCTTCGCGGTATGCTGTTTCG 192
|||||
Db 1 ACATGATGTCAGGTATCAAAAAGTCTGTGAAGTTCTTCGCGGTATGCTGTTTCG 60
|||||
QY 193 TTCCACAATAATAG 206
|||||
Db 61 TTCCACAATAATAG 74
|||||

RESULT 25
LOCUS AB003281 1128 bp DNA linear MAM 26-MAY-1999
DEFINITION Phacochoerus aethiopicus gene for elafin (trappin-2), partial cds.
ACCESSION AB003281
VERSION AB003281.1 GI:4887637
KEYWORDS elafin (trappin-2).
SOURCE Phacochoerus aethiopicus DNA.
ORGANISM Phacochoerus aethiopicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Phacochoerus.

```

REFERENCE 1 (sites)
 AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (bases 1 to 1128)
 AUTHORS Hirose,S.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 (E-mail:shirosebio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
 FEATURES
 source Location/Qualifiers
 1..1128
 /organism="Phacochoerus aethiopicus"
 /db_xref="taxon:85517"
 <211..632
 /codon_start=3
 /product="elafin (trappin-2)"
 /protein_id="BAA77825.1"
 /db_xref="GI:4887638"
 /translations="RPGKGTKEIHALVKQKVRGDEQVQGVFVKGDLGKQDPPV KAQLPDKQDLGKGEVSVKGDPIKAQLPDKQDVPVKAQPAIKRLILLTKPGSCPRIIL IRCMVNPPNRLSDAQCPGVKKCCGFCGKECLNPR"
 211..>632
 BASE COUNT 296 a 275 c 322 g 235 t
 ORIGIN
 Query Match 35.9%; Score 74; DB 4; Length 1128;
 Best Local Similarity 71.0%; Pred. No. 3e-12;
 Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 59 TACTAAGCCAGGTTCTGCTTATCTGATTCGTTGCGCTATGTTAAACCCACTAA 118
 Db 488 TACCAGCTGCTCTCTCCCGAGATTCTGATCCGTTGATGATGTCATCCCTCAA 547
 QY 119 CCGTGTGTTGAAGGACACTGATGTCAGGATCAAAAAGTGTGTAAGGTTCTCTCGG 178
 Db 548 CAGGTGTTGAGTGATGCTCAGTGCAGGCCAGGGTCAAGAAGTGTGTGAAGGCTTTTCGG 607
 QY 179 TATGCTTGTTCGTTCC 196
 Db 608 GAAGGAATGTTGAATCC 625
 RESULT 26
 PIGWAPD 464 bp mRNA linear MAM 14-APR-2000
 LOCUS Pig mRNA for elafin family member protein, complete cds.
 DEFINITION D50322
 ACCESSION D50322.1 GI:1228063
 VERSION elafin family member protein.
 KEYWORDS Sus scrofa cDNA to mRNA, clone:WAP-5.
 SOURCE Sus scrofa
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (sites)
 AUTHORS Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A., Tachibana,S. and Hirose,S.
 TITLE Accelerated evolution in inhibitor domains of porcine elafin family members
 JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
 MEDLINE 96215132
 REFERENCE 2 (sites)
 AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 3 (bases 1 to 464)
 AUTHORS Hirose,S.

Direct Submission
 Submitted (16-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 (E-mail:shirosebio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
 FEATURES
 Location/Qualifiers
 1..464
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone="WAP-5"
 <1..96
 /number=1
 primer_bind 1..19
 /note="PCR primer"
 5'UTR <1..14
 CDS 15..446
 /codon_start=1
 /product="elafin family member protein"
 /protein_id="BAA08857.1"
 /db_xref="GI:1228064"
 /translation="MRSRSFLVLVVFLICGTLVVOAGRIIRPKGKTKKTALVKQ QGPRGXDOVKQGPVKGDLGKSQDPVKAQLPDKQDVPVKAQPAIKRLILLTKPGSC PRIILCLMVNPPNRLSDAQCPGVKKCCGFCGKDCMDPK"
 15..77
 sig_peptide 261..443
 mat_peptide 97..447
 /product="unnamed"
 97..447
 /number=2
 447..464
 448..464
 /number=3
 primer_bind complement (448..464)
 /note="PCR primer"
 BASE COUNT 126 a 111 c 124 g 103 t
 ORIGIN
 Query Match 35.7%; Score 73.6; DB 4; Length 464;
 Best Local Similarity 69.4%; Pred. No. 4e-12;
 Matches 100; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
 QY 60 ACTAAGCCAGGTTCTGCTTATCTGATTCGTTGCGCTATGTTAAACCCACTAAC 119
 Db 303 ACCAAGCTGCTCTCTCCCGAGATTCTGATCCGTTGATGATGTCATCCCTCAA 362
 QY 120 CGTGTGTTGAAGGACACTGATGTCAGGATCAAAAAGTGTGTAAGGTTCTCTCGGT 179
 Db 363 AGGTGTTGAGTGATGCTCAGTGCAGGCCAGGGTCAAGAAGTGTGTGAAGGCTTTTCGGG 422
 QY 180 ATGGCTGTTGTTCCACAATAA 203
 Db 423 AAGGACTGTATGATCCCAAGTGA 446
 RESULT 27
 D83668 1034 bp DNA linear MAM 14-APR-2000
 LOCUS Sus scrofa gene for elafin homolog, exon2, partial cds.
 DEFINITION D83668
 ACCESSION D83668.1 GI:2055271
 VERSION elafin homolog; elastase inhibitor.
 KEYWORDS Sus scrofa DNA, clone_lib:PCR product.
 SOURCE Sus scrofa
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 1034)
 AUTHORS Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A., Tachibana,S. and Hirose,S.
 TITLE Accelerated evolution in the inhibitor domain of porcine elafin family members
 JOURNAL Journal of Biological Chemistry (1996) In press
 REFERENCE 2 (sites)
 AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and

Db 216 TGATCGGTGGCCATGTTGAATCCCTTAACCGTGTCTTGAAGATACGACTG 270

RESULT 29
A31081/c
LOCUS A31081 69 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI14 from patent EP0402068.
ACCESSION A31081
VERSION A31081.1 GI:1249295
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 69)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
JOURNAL against human elastase
PATENT: EP 0402068-A 21 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES Location/Qualifiers
source 1..69
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 26 a 12 c 16 g 15 t
ORIGIN

Query Match 33.5%; Score 69; DB 6; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 TTCTTGCTCTATTATCTGATTCGTTGGCTATGTTAAACCCACTAACCGTTGTTGAA 130
Db 69 TTCTTGCTCTATTATCTGATTCGTTGGCTATGTTAAACCCACTAACCGTTGTTGAA 10
Qy 131 GGACACTGA 139
Db 9 GGACACTGA 1

RESULT 30
A31080
LOCUS A31080 68 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI13 from patent EP0402068.
ACCESSION A31080
VERSION A31080.1 GI:1249294
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 68)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
JOURNAL against human elastase
PATENT: EP 0402068-A 20 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES Location/Qualifiers
source 1..68
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 13 a 16 c 14 g 25 t
ORIGIN

Query Match 33.0%; Score 68; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 GCCAGGTTCTTGCTATTATCTGATTCGTTGGCTATGTTAAACCCACTAACCGTTG 124
Db 1 GCCAGGTTCTTGCTATTATCTGATTCGTTGGCTATGTTAAACCCACTAACCGTTG 60
Qy 125 TTGGAAGG 132
Db 61 TTGGAAGG 68

RESULT 31
AB003283
LOCUS AB003283 1169 bp DNA linear MAM 26-MAY-1999
DEFINITION Pecari tajacu gene for trappin, partial cds.
ACCESSION AB003283
VERSION AB003283.1 GI:4887641
KEYWORDS trappin.
SOURCE Pecari tajacu DNA.
ORGANISM Pecari tajacu
REFERENCE 1 (sites)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Tayassuidae; Pecari.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1169)
AUTHORS Hirose,S.
DIRECT SUBMISSION
SUBMITTED (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail:shirosese@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
FEATURES Location/Qualifiers
source 1..1169
/organism="Pecari tajacu"
/db_xref="taxon:9829"
CDS <212..675
/codon_start=3
/product="trappin"
/protein_id="BAA77827.1"
/db_xref="GI:4887642"
/translation="RPRQGQTKGHALIKQDPPVRGQPVKQDPVKAQLPDKGQDLV KGQDPVKGQDLVKQDPPVKAQLPDKQDLPVKQDPPVKQDPPVKAELAVRRLV LPRKPGFCFPMKIRKALFNPRLCTDAGCPGARKCCIGSCGKACLNPR"

exon 305 a 289 c 330 g 245 t
BASE COUNT
ORIGIN

Query Match 32.9%; Score 67.8; DB 4; Length 1169;
Best Local Similarity 68.9%; Pred. No. 2.5e-10;
Matches 93; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 62 TAAGCCAGGTTCTTGCTCTATTATCTTGATTCGTTGGCTATGTTAAACCCACTAACCG 121
Db 531 TAAGCTGCTTCTGCCCATGATTAAAGATCGTTGTCCTGTTCAATCTCTCTAACAG 590
Qy 122 TTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTCGGTTAT 181
Db 591 GTGTTTGACCGATGCTGGTGCCAGGCCAGGAAGTGTCATAGGCTCTTTCGGGAA 650
Qy 182 GGCCTGTTTCGTTCC 196
Db 651 GGCCTGTTTGAATCC 665

RESULT 32
AB003284
LOCUS AB003284 734 bp DNA linear MAM 26-MAY-1999
DEFINITION Hippopotamus amphibius gene for trappin, partial cds.
ACCESSION AB003284
VERSION AB003284.1 GI:4887643
KEYWORDS trappin.
SOURCE Hippopotamus amphibius DNA.
ORGANISM Hippopotamus amphibius
REFERENCE 1 (sites)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1169)
AUTHORS Hirose,S.
DIRECT SUBMISSION
SUBMITTED (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail:shirosese@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)


```

Db 340 CTAAGCGTGGCCACTGCCCTAGGATCTCTTTTCGTGCGCGTGGCAATCCCTCTAACA 399
QY 121 GTTGTGTTGAAGGACACTGATTGCTCAGGTATCAAAAGTGTGTGAAGGTTCTCTCGGTA 180
Db 400 AGTGTGTGAGAGATTATGACTGTCCAGGGTCAAGAGTGTGTGAAGGCTTTTGCGGGA 459
QY 181 TGGCTGTTGTTGTTCCCAATAA 203
Db 460 AGGATTGTTGTATCCCAAGTGA 482

RESULT 35
LOCUS E06782
DEFINITION 737 bp RNA linear PAT 29-SEP-1997
ACCESSION cDNA containing a sequence encoding Na+,K+-ATPase inhibitory
VERSION peptide (SPAI:sodium ion, potassium ion ATPase inhibitory peptide).
E06782
KEYWORDS E06782.1 GI:2174964
SOURCE JP 1994049098-A/1.
ORGANISM Sus sp.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
TITLE Kuroki,A., Kuwata,M. and Tachibana,S.
JOURNAL NA+, K+-ATPASE INHIBITING AND DNA CODING THE SAME
EISAI CO LTD
COMMENT OS Sus sp. (pig)
PN JP 1994049098-A/1
PD 22-FEB-1994
PI 29-JUL-1992 JP 1992202286
PF KUROKI ATSUSHI, KUWATA MANABU, TACHIBANA SHINRO PC
C07K13/00,A61K37/64,A61K49/00,C12N15/12;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=Duodenum;
CC *source: library=lambda gt10 cDNA library;
CC *source: clone=PD-57G;
FH Key Location/Qualifiers
FT CDS 1..570
FT /product='SPAI'
FT mat_peptide 385..567
FT /product='SPAI'
FT misc_feature 70..567
FT /note='The region corresponds to high FT
FT 3'UTR molecular SPAI',
FT Location/Qualifiers
FT 1..737
FT /organism='Sus sp.'
FT /db_xref='taxon:9826'

BASE COUNT 205 a 192 c 182 g 158 t
ORIGIN

Query Match 32.1%; Score 66.2; DB 6; Length 737;
Best Local Similarity 66.4%; Pred. No. 8e-10;
Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 61 CTAAGCAGGTTCTGTCTATTATCTTGGTTCGTTGGCTATGTTAAACCCACCTAAC 120
Db 428 CTAAGCGTGGCCACTGCCCTAGGATCTCTTTTCGTGCGCGTGGCAATCCCTCTAACA 487
QY 121 GTTGTGTTGAAGGACACTGATTGCTCAGGTATCAAAAGTGTGTGAAGGTTCTCTCGGTA 180
Db 488 AGTGTGTGAGAGATTATGACTGTCCAGGGTCAAGAGTGTGTGAAGGCTTTTGCGGGA 547
QY 181 TGGCTGTTGTTGTTCCCAATAA 203

```

```

Db 548 AGGATTGTTGTATCCCAAGTGA 570

RESULT 36
LOCUS DMV245
DEFINITION Sus scrofa mRNA for preproSPAI-2, complete cds.
ACCESSION D83667 D17753
VERSION D83667.1 GI:1304046
KEYWORDS APAI-2; preproSPAI-2; proSPAI-2; SPAI-2.
SOURCE Sus scrofa cell_line:NM514 cDNA to mRNA, clone_lib:lambda gt10.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (sites)
AUTHORS Kuroki,J., Hosoya,T., Itakura,M., Hirose,S., Tamechika,I.,
Yoshimoto,T., Ghoneim,M.A., Nara,K., Kato,A., Suzuki,Y.,
Furukawa,M. and Tachibana,S.
TITLE Cloning, characterization, and tissue distribution of porcine SPAI,
a protein with a transglutaminase substrate domain and the WAP
motif
JOURNAL J. Biol. Chem. 270 (38), 22428-22433 (1995)
MEDLINE 95403443
REFERENCE 2 (bases 1 to 789)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; Ookayama, Meguro-ku,
Tokyo 152, Japan (Tel:81-3-3726-1111, Fax:81-3-3729-0335)
On or before Mar 17, 1999 this sequence version replaced
gi:2078450, gi:1054611, gi:2077950.
COMMENT D17753:Submitted (27-Sep-1993) to DDBJ by:Shigehisa Hirose.
FEATURES
source
location/Qualifiers
1..789
/organism='Sus scrofa'
/db_xref='taxon:9823'
/cell_line='NM514'
/clone_lib='lambda gt10'
61..624
/codon_start=1
/product='preproSPAI-2'
/protein_id='BAA12037.1'
/db_xref='GI:1304181'
/translation='MRSRSFLVLVAVFLICETLVAORLDRIRPKGQGPVEGQDQD
EGPGFVKVEILDIGDQFVKQDPVKQDPVKQDLPVKQDLPVKQDLPVKQDLPVKQDLPVKQD
VVGHEPVEGQDPVNAQLPKQDPVKQDPVKQDPVKQDPVKQDPVKQDPVKQDPVKQDPVKQD
KCWRDYDCPGVKKCEGFCGKDCLYPK'
61..123
sig_peptide
mat_peptide 124..621
/product='proSPAI-2'
mat_peptide 439..621
/product='SPAI-2'
polyA_site 722
polyA_signal 761..766
BASE COUNT 201 a 215 c 197 g 176 t
ORIGIN

Query Match 32.1%; Score 66.2; DB 4; Length 789;
Best Local Similarity 66.4%; Pred. No. 7.9e-10;
Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 61 CTAAGCAGGTTCTGTCTATTATCTTGGTTCGTTGGCTATGTTAAACCCACCTAAC 120
Db 482 CTAAGCGTGGCCACTGCCCTAGGATCTCTTTTCGTGCGCGTGGCAATCCCTCTAACA 541
QY 121 GTTGTGTTGAAGGACACTGATTGCTCAGGTATCAAAAGTGTGTGAAGGTTCTCTCGGTA 180
Db 542 AGTGTGTGAGAGATTATGACTGTCCAGGGTCAAGAGTGTGTGAAGGCTTTTGCGGGA 601
QY 181 TGGCTGTTGTTGTTCCCAATAA 203
Db 602 AGGATTGTTGTATCCCAAGTGA 624

```

```
RESULT 37
PIGWAPP
LOCUS
DEFINITION Pig DNA for SPAI-2, complete cds.
ACCESSION D50320
VERSION D50320.1 GI:1228059
KEYWORDS SPAI-2.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-2.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
    1 (sites)
    Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
    Tachibana, S. and Hirose, S.
    Accelerated evolution in inhibitor domains of porcine elafin family
    members
    J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3782)
AUTHORS Hirose, S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3782)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
    Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
    Midori-ku, Yokohama, Kanagawa 226-8501, Japan
    (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
    Fax: 045-924-5824)
FEATURES
    source
        1..3782
        /organism="Sus scrofa"
        /db_xref="taxon:9823"
        /clone="lambda WAP-2"
        /clone_lib="EMBL SP6/T7"
        <1..710
        /number=1
        /evidence=experimental
        join(629..710,1552..2033)
        /codon_start=1
        /evidence=experimental
        /product="SPAI-2"
        /protein_id="BAA08855.1"
        /db_xref="GI:1228060"
        /translation="WRSRSLVFLVAVFLICETLVAQLDRIRGPKGQDPVEGQDQD
        EGPVPVKVILIDIQDPVKGQDPVKGQDPVKGQDLVKSDPVPKAEIPDQIGD
        VKGHEPVEGQDPVNAQLPDKVQDPVKAQPAVPGRPFLLSKRGHCPRIILFRCLNPSN
        KCRWDYDCPVGKCCGEGFCGKDCLYPK"
        629..691
        /evidence=experimental
        1848..2030
        /product="SPAI-2"
        /evidence=experimental
        711..1551
        /number=1
        /evidence=experimental
        1552..2034
        /number=2
        /evidence=experimental
        join(2034,2545..2701)
        2035..2544
        /number=2
        /evidence=experimental
        2240..2502
        /notes="SINE (PRE-1)"
        2545..2701
        /number=3
        /evidence=experimental
        2680..2685
        /evidence=experimental
        2701
        /evidence=experimental
    sig_peptide
    mat_peptide
    intron
    exon
    3'UTR
    intron
    repeat_unit
    exon
    polyA_signal
    polyA_site
```

```
repeat_region 3671..3742
/notes="TG repeat"
BASE COUNT 823 a 972 c 1149 g 838 t
ORIGIN
Query Match 32.1%; Score 66.2; DB 4; Length 3782;
Best Local Similarity 66.4%; Pred. No. 7.7e-10; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 48;
Qy 61 CTAAGCCAGGTTCTGTGCTTATTATCTTGGTTCGCTATGTTAAACCCACCTAAC 120
    |||||
Db 1891 CTAAGCGTGCCACTGCCCTAGGATTCCTTTGTTGCCCGCTGAGCAATCCCTCTAAC 1950
Qy 121 GTTCTTTGAAGGACACTGATGTGTCAGGATTCAAAAAGTGTGTGAAGGTTCTTCGCGGTA 180
    |||||
Db 1951 AGTGTGTGAGAGATTATGACTGTCCAGGGGTCAAGAGTGTGTGAAGGCTTTTTCGCGGA 2010
Qy 181 TGGCTTCTTTCGTTCCACATAA 203
    |||||
Db 2011 AGGATTGTTGTATCCCAAGTGA 2033
RESULT 38
A31079/c
LOCUS A31079
DEFINITION Oligonucleotide ELI2 from patent EP0402068.
ACCESSION A31079
VERSION A31079.1 GI:1249293
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 66)
AUTHORS Christophers, E., Schroder, J.M., Pioli, D., Wiedow, O. and Edge, M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
    against human elastase
    Patent: EP 0402068-A 19 12-DEC-1990;
    IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
    source
        1..66
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        18 t
BASE COUNT 15 a 15 c 18 g 18 t
ORIGIN
Query Match 32.0%; Score 66; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 9.6e-10; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 0;
Qy 5 CGAGTCGTGTACCATCTGTCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAA 64
    |||||
Db 66 CGAGTCGTGTACCATCTGTCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAA 7
Qy 65 GCCAGG 70
    |||||
Db 6 GCCAGG 1
RESULT 39
A31078
LOCUS A31078
DEFINITION Oligonucleotide ELI1 from patent EP0402068.
ACCESSION A31078
VERSION A31078.1 GI:1249292
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 64)
AUTHORS Christophers, E., Schroder, J.M., Pioli, D., Wiedow, O. and Edge, M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
    against human elastase
    Patent: EP 0402068-A 18 12-DEC-1990;
JOURNAL
```

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES

source

1..64

Location/Qualifiers

/organism="synthetic construct"

/db_xref="taxon:32630"

19 a 16 c 12 g 17 t

BASE COUNT

ORIGIN

Query Match 31.1%; Score 64; DB 6; Length 64;
 Best Local Similarity 100.0%; Pred. No. 4e-09;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGTACCATCTGATATGCTCAAGAACCAAGTTAAAGTCTGTGCTA 60
 Db 1 AATTCGAGCTCGTACCATCTGATATGCTCAAGAACCAAGTTAAAGTCTGTGCTA 60

QY 61 CTA 64

Db 61 CTA 64

RESULT 40
 AB003282
 LOCUS 1325 bp DNA linear MAM 26-MAY-1999
 DEFINITION Phacochoerus aethiopicus gene for SPAI (trappin-1), partial cds.
 ACCESSION AB003282
 VERSION SPAI (trappin-1).
 KEYWORDS Phacochoerus aethiopicus DNA.
 SOURCE Phacochoerus aethiopicus
 ORGANISM Phacochoerus aethiopicus

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 2 (sites)

Db 686 CTAAGCGTGGCGCTGCCCTTGGATTCTTCTCGTTCGCCCGCAATCCTCTAACA 745

QY 121 GTTGTGTTGAAGACACTGATTGTCCAGGTATCAAAAAGTGTGAAGCTTCTCGGTA 180

Db 746 AGTGTGGAGAGATTAAGTACTGTCCAGGGTCAAGAAGTGTGTGAAGCTTTTCGGGA 805

QY 181 TGGCTTGTGTTTCCACATAA 203

Db 806 AGGATTGTTGTATCCCAAGTGA 828

RESULT 41
 PIGWAPE 578 bp mRNA linear MAM 14-APR-2000
 LOCUS Pig mRNA for elafin family member protein, complete cds.
 DEFINITION D50323.1 GI:1228065
 ACCESSION elafin family member protein.
 VERSION Sus scrofa cDNA to mRNA, clone:WAP-4.
 KEYWORDS Sus scrofa
 SOURCE Sus scrofa
 ORGANISM Sus scrofa

REFERENCE
 AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 TITLE Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A., Tachibana, S. and Hirose, S.
 JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
 MEDLINE 96215132
 REFERENCE 2 (sites)

REFERENCE
 AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
 TITLE Evolution of the trappin multigene family in the Suidae
 JOURNAL J. Biochem. 124 (3), 491-502 (1998)
 MEDLINE 98391820
 REFERENCE 3 (bases 1 to 578)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
 AUTHORS Hirose, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
 MEDLINE (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)


```

BASE COUNT      169 a      141 c      138 g      130 t
ORIGIN
Query Match      29.3%; Score 60.4; DB 4; Length 578;
Best Local Similarity 64.1%; Pred. No. 5.1e-08;
Matches 91; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 62 TAAGCCAGGTTCTGTCCTATATCTTGTGATTCGTTGCGTGTGTAACCCACCTAACCG 121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 TAAGCCTGGCCTCTGCCCTTGGATTTTACGTTGCCGCTGCCCAACCTCCTAACAA 478
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 122 TTGTTTGAAGGACACTGATGTCAGGTATCAAAAAGTCTGTGAAGGTTCTCTGGCGTAT 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 ATGTTGAGAGATTCCTACTGCCAGGGTCAATGAAGTCTGTGAAGGTTTTTGGCGGAA 538
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 182 GGCTTGTGTTCCGTTCCACAATAA 203
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 TGAATGTTCTATCCCGGATGA 560
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 42
AB011010 AB011010 495 bp DNA linear MAM 14-APR-2000
LOCUS Bos taurus gene for Trappin-6, partial cds.
DEFINITION AB011010
ACCESSION AB011010
VERSION AB011010.1 GI:3132273
KEYWORDS Trappin-6.
SOURCE Bos taurus DNA.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (sites)
AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 495)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1998) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail: shirose@bio.titech.ac.jp, Tel: +81-45-924-5726, Fax: +81-45-924-5824)
FEATURES
source 1..495
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..36
/number=1
primer_bind 1..24
/notes="PCR primer site"
<37..329
CDS /codon_start=3
/product="Trappin-6"
/protein_id="BAA28148.1"
/db_xref="GI:3132274"
/translation="SPKGGNVVFNKGFPVNGQSPDKGQDPVKGGQDPVKGGQVVVAQD
RAGLPFRGLCPVRVIRHCLNPPNQWRDAHCPGAKKCCGFCGKTCMNR"
37..329
exon 37..495
primer_bind 472..495
/notes="PCR primer site"
BASE COUNT      134 a      99 c      163 g      99 t
ORIGIN
Query Match      26.5%; Score 54.6; DB 4; Length 495;
Best Local Similarity 61.7%; Pred. No. 3.3e-06;
Matches 87; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 63 AAGCAGGTTCTGCTCTATATCTTGTGATTCGTTGCGCTATGTAACCCACCTAACCGT 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 189 AAGCGTGGCTTATGCCCCAGGGTTCCGATCCACTGCAACTTGTGGAAATCCCCCTAACACAG 248
QY 123 TGTTTGAAGGACACTGATGTTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCTGGCGTATG 182
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 TGTGGAGAGATGCTACTGCCAGGGGCCAAGAGTGTGTGAAGGCTTTTGTGGGAAG 308
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 GCTTGTGTTCCGTTCCACAATAA 203
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 ACCTGTATGAATCCCGGTGA 329
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 43
PIGWAPC
LOCUS Pig DNA for elafin family member protein, complete cds.
DEFINITION D50321
ACCESSION D50321
VERSION D50321.1 GI:1228061
KEYWORDS elafin family member protein.
SOURCE Sus scrofa
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (sites)
AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A., Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3670)
AUTHORS Hirose, S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3670)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)
FEATURES
source 1..3670
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="lambda WAP-3"
/clone_lib="EMBL SP6/T7"
<1..710
/number=1
5'UTR join(629..710,1552..1904)
CDS /codon_start=1
/product="elafin family member protein"
/protein_id="BAA08856.1"
/db_xref="GI:1228062"
/translation="MRSRSLVAVFLICETLVAQRLRIRGPKGGQDPVGGQDQD
EGQGPVKVEILDQDLVKGQDPVGAQLPKVDQDPVKAQPIQGGLFPRPG
VCPKIFCPVLNPPPIKWRDSDHCPGVKKCKGKCVTPR"
629..691
sig_peptide 1722..1901
mat_peptide /product="unnamed"
intron 711..1551
/number=1
exon 1552..1905
/number=2
3'UTR join(1905,2414..2570)
intron 1906..2413
/number=2
repeat_unit 2110..2371
/notes="SINE (PRE-1)"
exon 2414..2570
/number=3
polyA_signal 2549..2554
polyA_site 2570

```

181 TG

* 14609 18981: contig of 4373 bp in length
* 18982 19081: gap of unknown length
* 19082 22382: contig of 3301 bp in length
* 22383 22482: gap of unknown length
* 22483 24761: contig of 2279 bp in length
* 24762 24861: gap of unknown length
* 24862 28297: contig of 3436 bp in length
* 28298 31802: contig of 3405 bp in length
* 31803 31902: gap of unknown length
* 31903 37457: contig of 5555 bp in length
* 37458 37557: gap of unknown length
* 37558 43403: contig of 5846 bp in length
* 43404 43503: gap of unknown length
* 43504 49186: contig of 5683 bp in length
* 49187 49286: gap of unknown length
* 49287 56048: contig of 6762 bp in length
* 56049 56148: gap of unknown length
* 56149 63736: contig of 7588 bp in length
* 63737 63836: gap of unknown length
* 63837 69493: contig of 5657 bp in length
* 69494 69593: gap of unknown length
* 69594 77974: contig of 8381 bp in length
* 77975 78074: gap of unknown length
* 78075 83932: contig of 5858 bp in length
* 83933 84032: gap of unknown length
* 84033 92690: contig of 8658 bp in length
* 92691 92790: gap of unknown length
* 92791 103073: contig of 10283 bp in length
* 103074 103173: gap of unknown length
* 103174 114922: contig of 11749 bp in length
* 114923 115022: gap of unknown length
* 115023 127345: contig of 12323 bp in length
* 127346 127445: gap of unknown length
* 127446 138012: contig of 10567 bp in length
* 138013 138112: gap of unknown length
* 138113 148920: contig of 10808 bp in length
* 148921 149020: gap of unknown length
* 149021 164577: contig of 15557 bp in length
* 164578 164678: gap of unknown length
* 164679 180799: contig of 16122 bp in length
* 180800 180899: gap of unknown length
* 180900 204225: contig of 23326 bp in length.

FEATURES
Location/Qualifiers
1..204225
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-170K14"

BASE COUNT 61390 a 37556 c 38806 g 58187 t 8286 others
ORIGIN

Query Match 17.3%; Score 35.6; DB 2; Length 204225;
Best Local Similarity 54.6%; Pred. No. 2.4;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 76 GTCTATTATCTTGATCTCGGTGATGTTAAACCCACCGTGTGTTGAGGACA 135
Db 109312 GTCCTGGTCTCTCTCTCCAAACAGTGAATCTCTCAATCTTGAGTCTGTCC 109371

QY 136 CTGATTGTCAGGATCAAAAAGTGTGTAAGGTTCTCGCGTATGCGTGTGTTTCGTTTC 195
Db 109372 CCTTTCCTCAGGACCTAAAGTTCTGCTATTCTCTGCGCTATGCTTCTTACTGA 109431

QY 196 CACAATAATA 205
Db 109432 CAGACCAAGA 109441

RESULT 48
AC006054/c 143738 bp DNA linear PRI 07-MAR-2002
LOCUS Homo sapiens Xq28 BAC GSI-382P7 (Roswell Park Cancer Institute
DEFINITION Human BAC Library) complete sequence.

AC006054
AC006054.2 GI:4557048
HTG.
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143738)
Muzny,D., Aronson,A.D., Bouck,J., Brundage,E., Bunac,C., Chen,Z.,
Di,W., Ding,Y., Dugan,S., Durbin,J., Forcum,J., Garcia,C.,
Gorell,J.H., Gorrell,L.L., Hernandez,J., Jackson,L.,
Kondejewski,N., Leal,B., Lichtarge,O., Liu,W., Logan,O., Lu,J.,
Martinez,C., Oswal,G., Pampell,L.R., Parish,B.J., Perez,L.,
Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Simon,M., Vo,Q.,
Williamson,A., Worley,K.C., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (26-NOV-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (31-MAR-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (01-APR-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (28-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 31, 1999 this sequence version replaced gi:4225895.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2


```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 146935)
JOURNAL Sultston,J.E. and Waterston,R.
MEDLINE Toward a complete human genome sequence
PUBMED Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 146935)
AUTHORS Nguyen,C., Abbott,A., Dignan,G., Doebber,A. and Boyer,E.
TITLE The sequence of Homo sapiens BAC clone RP11-2G16
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 146935)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2001) Genome Sequencing Center, Washington
MO 63108, USA
4 (bases 1 to 146935)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington
UNIVERSITY, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Sep 19, 2001 this sequence version replaced gi:8072447.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
-----
Summary Statistics
Center project name: H_NH0002G16
Drafting Center: WIBR
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Iateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-403L5. Actual start of
this clone is at base position 1 of RP11-2G16; actual end is at
base position 146935 of RP11-2G16.

Data from AC078976 was used to finish this clone, AC021047.

The sequence of AC021047 has been incorporated into AC096647.
Location/Qualifiers

```

```

repeat_region 25582..25877
/rpt_family="Alu"
repeat_region 25863..25922
/rpt_family="AT_rich"
repeat_region 25895..26547
/rpt_family="L1"
repeat_region 26552..26649
/rpt_family="L1"
repeat_region 26649..27752
/rpt_family="L1"
repeat_region 27753..28143
/rpt_family="MaLR"
repeat_region 28150..28424
/rpt_family="MaLR"
repeat_region 28425..28453
/rpt_family="T"n"
repeat_region 28605..29988
/rpt_family="MaLR"
repeat_region 30035..30248
/rpt_family="Alu"
repeat_region 30260..30999
/rpt_family="L1"
repeat_region 31000..31073
/rpt_family="MaLR"
repeat_region 31079..31311
/rpt_family="Alu"
repeat_region 31330..31623
/rpt_family="ERV1"
repeat_region 31648..31712
/rpt_family="L1"
repeat_region 31720..33118
/rpt_family="L1"
repeat_region 33111..33144
/rpt_family="AT_rich"
repeat_region 33125..33921
/rpt_family="L1"
repeat_region 35401..38116
/rpt_family="L1"
repeat_region 35601..35634
/rpt_family="AT_rich"
repeat_region 37164..37213
/rpt_family="T-rich"
repeat_region 38095..38245
/rpt_family="L1"
repeat_region 38267..39514
/rpt_family="MaLR"
repeat_region 39515..39924
/rpt_family="MaLR"

Query Match 17.1%; Score 35.2; DB 9; Length 146935;
Best Local Similarity 54.7%; Pred. No. 3.2;
Matches 70; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 44 TAAAGTCCTGCTACTAAGCAGGTCCTTCTCTATATCTTGAATTCGTTGCGCTAT 103
| | | | |
Db 31575 TAAAGATGAGCTTTTTCACAGGTTATGCTTTTATGATAGTTGTTGTTGCAATTT 31634
| | | | |

QY 104 GTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAGGATCAAAAAGTGCTG 163
| | | | |
Db 31635 TTTAACTAAACAATAATTGTTTATGTCATATAATGTTAGTATCAGATCAGGTTG 31694
| | | | |

QY 164 TGAAGGTT 171
| | | | |
Db 31695 TGTAGGT 31702

RESULT 50
PIGALP PIGALP
LOCUS Porcine antileukoproteinase mRNA, complete cds.
DEFINITION Porcine antileukoproteinase mRNA, complete cds.
ACCESSION M57446
VERSION M57446.1 GI:164319
KEYWORDS antileukoproteinase.

SOURCE Porcine uterus. cDNA to mRNA, clone pALP.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 600)
AUTHORS Farmer, S.J., Fliss, A.E. and Simmen, R.C.
TITLE Complementary DNA cloning and regulation of expression of the
messenger RNA encoding a pregnancy-associated porcine uterine
protein related to human antileukoproteinase
JOURNAL Mol. Endocrinol. 4 (8), 1095-1104 (1990)
MEDLINE 91155942
PUBMED 2293019

FEATURES
source Location/Qualifiers
1..600
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="pALP"
/sex="female"
/tissue_type="uterus"
1..600
/gene="ALP"
44..391
/gene="ALP"
/codon_start=1
/product="antileukoproteinase"
/protein_id="AA63446.1"
/db_xref="GI:164320"
/translation="MAPWAVEGAENALKGGACPPRKIVOCRLRYEKPKCTSDWQCPDKK
KCRDTCATKCLNPVAITNPVKPGKVPVYQGCWMLNPPNHCKTDSQCLGDLKCKK
SMCGKVCLTPVKA"
600
polyA_site
BASE COUNT 146 a 156 c 151 g 147 t
ORIGIN
Query Match 16.9%; Score 34.8; DB 4; Length 600;
Best Local Similarity 53.7%; Pred. No. 4.7;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
| | | | |
Db 245 AAGCTGGGAAGTGTCCAGTGGTCTATGCGCAGTGATGATGCTCAACCCCAATCAC 304
| | | | |

QY 123 TGTTTGAAGGACACTGATTGTCAGGTATCAAAAGTGCTGTGAAGTTCTCTGGGTATG 182
| | | | |
Db 305 TGCAAGACAGACAGCCAGTGCCTGGTGACTTAAATGCTCAAGAGCATGTGCGGAAA 364
| | | | |

QY 183 GCTTGTTCGTTCC 196
| | | | |
Db 365 GTCTGCCTCACCCC 378

RESULT 51
MMU808093
LOCUS Mus musculus secretory leukocyte protease inhibitor mRNA, complete
DEFINITION cds.
ACCESSION U88093
VERSION U88093.1 GI:1945382
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 671)
AUTHORS Zitnik, R.J., Zhang, J., Kashem, M.A., Kohno, T., Lyons, D.E.,
Wright, C.D., Rosen, E., Goldberg, I. and Hayday, A.C.
TITLE The cloning and characterization of a murine secretory leukocyte
protease inhibitor cDNA
JOURNAL Biochem. Biophys. Res. Commun. 232 (3), 687-697 (1997)
MEDLINE 97271386
PUBMED 9126337
REFERENCE 2 (bases 1 to 671)

```

AUTHORS Zitnik, R.J., Zhang, J., Kashem, M.A., Kohno, T., Lyons, D.E., Wright, C.D., Rosen, E., Goldberg, I. and Hayday, A.C.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-1997) Internal Medicine, Yale University, 333 Cedar Street, New Haven, CT 06520, USA
 FEATURES Location/Qualifiers
 source 1..671
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="ras-transformed 3T3 cells"
 CDS 13..408
 /function="serine protease inhibitor"
 /codon_start=1
 /product="secretory leukocyte protease inhibitor"
 /protein_id="AAC53140.1"
 /db_xref="GI:1945383"
 /translation="MKSCGLLPFTVLALGILAPWTVEGGKNDIKIGACPAKPKPAQC LKLEKPCQRTDWECPGKRCODACGSKCNPVPIRKPVRKPGRCVKVQARCMWLNPNVQQRGQCDGKYKCCGICGKVCCLPPM"
 BASE COUNT 157 a 179 g 177 g 158 t
 ORIGIN
 Query Match 16.8%; Score 34.6; DB 10; Length 671;
 Best Local Similarity 53.3%; Pred. No. 5.4;
 Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 51 CCTGTGCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGTGTTGGCTATGTTAAAC 110
 |||||
 Db 256 CCAGTGTGGAGGAGCGCTGGAGGTGCGTCAAAACTCAGGCAAGATGATGATGCTTAAAC 315
 |||||
 QY 111 CCACCTAACCGTTGTTTGAAGGACACTGATGTCAGGTATCAAAAGTCTGTGAAGGT 170
 |||||
 Db 316 CTTCCCAATGCTGCGCAGGAGCGGCGAGTGTGACGGCAAAATCAAGTCTGTGAGGGT 375
 |||||
 QY 171 TCTCGCGTATGGCTTG 187
 |||||
 Db 376 ATATGTGGGAAAGTCTG 392
 |||||
 RESULT 52
 MMU94341
 LOCUS Mus musculus secretory leukoprotease inhibitor mRNA, complete cds.
 DEFINITION
 ACCESSION U94341
 VERSION U94341.1 GI:1945450
 KEYWORDS
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 682)
 AUTHORS Abe, T., Tomimaga, Y., Kikuchi, T., Watanabe, A., Satoh, K., Watanabe, Y. and Nukiwa, T.
 TITLE Bacterial pneumonia causes augmented expression of the secretory leukoprotease inhibitor gene in the murine lung
 JOURNAL Am. J. Respir. Crit. Care Med. 156 (4 Pt 1), 1235-1240 (1997)
 MEDLINE 9351627
 PUBMED
 REFERENCE 2 (bases 1 to 682)
 AUTHORS Abe, T., Tomimaga, Y., Kikuchi, T., Watanabe, A., Satoh, K., Watanabe, Y. and Nukiwa, T.
 TITLE Direct Submission
 JOURNAL Submitted (18-MAR-1997) Department of Respiratory Oncology and Molecular Medicine, Institute of Development, Aging and Cancer, Tohoku University, 4-1 Seiryomachi, Aoba-Ku, Sendai 980-77, Japan
 FEATURES Location/Qualifiers
 source 1..682
 /organism="Mus musculus"
 /strain="B6CBA [C57Black/6 x CBA]"
 /db_xref="taxon:10090"
 /tissue type="lung"
 CDS 14..409
 /codon_start=1

/product="secretory leukoprotease inhibitor"
 /protein_id="AAC53394.1"
 /db_xref="GI:1945451"
 /translation="MKSCGLLPFTVLALGILAPWTVEGGKNDIKIGACPAKPKPAQC LKLEKPCQRTDWECPGKRCODACGSKCNPVPIRKPVRKPGRCVKVQARCMWLNPNVQQRGQCDGKYKCCGICGKVCCLPPM"
 BASE COUNT 163 a 179 c 181 g 159 t
 ORIGIN
 Query Match 16.8%; Score 34.6; DB 10; Length 682;
 Best Local Similarity 53.3%; Pred. No. 5.4;
 Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 51 CCTGTGCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGTGTTGGCTATGTTAAAC 110
 |||||
 Db 257 CCAGTGTGGAGGAGCGCTGGAGGTGCGTCAAAACTCAGGCAAGATGATGATGCTTAAAC 316
 |||||
 QY 111 CCACCTAACCGTTGTTTGAAGGACACTGATGTCAGGTATCAAAAGTCTGTGAAGGT 170
 |||||
 Db 317 CTTCCCAATGCTGCGCAGGAGCGGCGAGTGTGACGGCAAAATCAAGTCTGTGAGGGT 376
 |||||
 QY 171 TCTCGCGTATGGCTTG 187
 |||||
 Db 377 ATATGTGGGAAAGTCTG 393
 |||||
 RESULT 53
 BC028509
 LOCUS Mus musculus secretory leukocyte protease inhibitor, clone
 DEFINITION MGC:41142 IMAGE:1513866, mRNA, complete cds.
 ACCESSION BC028509
 VERSION BC028509.1 GI:20306995
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 894)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Marcello Bento Soares, Ph.D.
 CDNA Library Preparation: Soares Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Brin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Mitanada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 67 Row: j Column: 6
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6755573.
 FEATURES Location/Qualifiers
 source 1..894
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:41142 IMAGE:1513866"


```

/tissue_type="Mammary gland, lactating mouse"
/clone_lib="Soares_mammary_gland_NMLMG"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac"
236..631
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAH28509.1"
/db_xref="GI:20306996"
/db_xref="LocusID:20568"
/translation="MKSCGLLPFTVLALGILAPWTVEGKNDAIKIGACPAKPAQC
LKLEKPCQRTDWECPGKQRCODAGSKCVNVPKRPVWRKRCVKTQARCMNLNP
PNCVQDQCGDKYKCCGICGKVLPPM"
BASE COUNT      205 a 242 c 233 g 214 t
ORIGIN
Query Match      16.8%; Score 34.6; DB 10; Length 894;
Best Local Similarity 53.3%; Pred. No. 5.4;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 51 CCTGTGCTACTAAGCCAGGCTTCTGCTCTATTATCTTGATTGCTTGGCGCTATGTTAAAC 110
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 CAGTGTGAGGAGCTGGAGGTGCTGCAAACTCAGGCAAGATGATGATGCTTAAC 538
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 111 CCACCTAACCGTTGTTTGAAGGACACTGATGTCTCCAGGTATCAAAAAGTCTGTGAAGGT 170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 CTTCCCAATGCTCTGCAGAGGACGGCAGTGTGACGGCAATACAAGTGTGTGAGGCT 598
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 171 TCTCGGGTATGCTTG 187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 ATATGTGGGAAGTCTG 615

RESULT 54
MMU73004
LOCUS
DEFINITION
Mus musculus secretory leukocyte protease inhibitor mRNA, complete cds.
U73004.1 GI:1763262
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1123)
Jin,F.Y., Nathan,C., Radzioch,D. and Ding,A.
Secretory leukocyte protease inhibitor: a macrophage product
induced by and antagonistic to bacterial lipopolysaccharide
Cell 88 (3), 417-426 (1997)
97191310
PUBMED
9039268
REFERENCE
2 (bases 1 to 1123)
Ding,A., Jin,F.-Y. and Nathan,C.F.
Direct Submission
Submitted (01-OCT-1996) Medicine, Cornell University Medical
College, 1300 York Ave. Box 57, New York, NY 10021, USA
FEATURES
source
1..1123
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="RAW 264.7"
/cell_type="macrophage"
447..842
/function="serine protease inhibitor"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAC53047.1"
/db_xref="GI:1763263"
/translation="MKSCGLLPFTVLALGILAPWTVEGKNDAIKIGACPAKPAQC
LKLEKPCQRTDWECPGKQRCODAGSKCVNVPKRPVWRKRCVKTQARCMNLNP
PNCVQDQCGDKYKCCGICGKVLPPM"
BASE COUNT      286 a 297 c 288 g 252 t

```

```

ORIGIN
Query Match      16.8%; Score 34.6; DB 10; Length 1123;
Best Local Similarity 53.3%; Pred. No. 5.4;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 51 CCTGTGCTACTAAGCCAGGCTTCTGCTCTATTATCTTGATTGCTTGGCGCTATGTTAAAC 110
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 CCAAGTGTGGAGGAAGCTGGGAGGTGCGTCAAACTCAGGCAAGATGATGATGCTTAAC 749
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 111 CCACCTAACCGTTGTTTGAAGGACACTGATGTCTCCAGGTATCAAAAAGTCTGTGAAGGT 170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 750 CTTCCCAATGCTCTGCAGAGGACGGCAGTGTGACGGCAATACAAGTGTGTGAGGCT 809
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 171 TCTCGCGTATGCTTG 187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 810 ATATGTGGGAAGTCTG 826

RESULT 55
AC110686
LOCUS
DEFINITION
Rattus norvegicus clone CH230-197B1, *** SEQUENCING IN PROGRESS
***, 81 unordered pieces.
AC110686
AC110686.3 GI:21739307
HTG; HTGS PHASE1.
KEYWORDS
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 156460)
REFERENCE
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaral, H.C., Are,J.R., Avele,M., Banks,T.,
Barbieri,J., Benton,J., Bimberg,K., Blankenburg,K., Bonnin,D.,
Bouck,J.E., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J.J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Prantze,P.,
Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,A., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okwundu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Uman,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Direct Submission
 Unpublished
 2 (bases 1 to 156460)
 Worley, K.C.
 Direct Submission
 Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 156460)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18767289.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRKM
 Center Clone name: CH230-197B1
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 89134 bases at least Q40
 Consensus quality: 94194 bases at least Q30
 Consensus quality: 97800 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Cenbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 81 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

20567	22031:	contig of 1465 bp in length
22032	22131:	gap of unknown length
22132	23615:	contig of 1484 bp in length
23616	23715:	gap of unknown length
23716	25016:	contig of 1301 bp in length
25017	25116:	gap of unknown length
25117	26162:	contig of 1046 bp in length
26163	26262:	gap of unknown length
26263	27694:	contig of 1432 bp in length
27695	27794:	gap of unknown length
27795	29680:	contig of 1886 bp in length
29681	29780:	gap of unknown length
29781	30780:	contig of 1000 bp in length
30781	30880:	gap of unknown length
30881	31905:	contig of 1025 bp in length
31906	32005:	gap of unknown length
32006	33691:	contig of 1686 bp in length
33692	33791:	gap of unknown length
33792	35109:	contig of 1318 bp in length
35110	35209:	gap of unknown length
35210	36589:	contig of 1380 bp in length
36590	36689:	gap of unknown length
36690	38294:	contig of 1605 bp in length
38295	38394:	gap of unknown length
38395	40181:	contig of 1787 bp in length
40182	40281:	gap of unknown length
40282	41573:	contig of 1292 bp in length
41574	41673:	gap of unknown length
41674	43896:	contig of 2223 bp in length
43897	43996:	gap of unknown length
43997	45168:	contig of 1172 bp in length
45169	45268:	gap of unknown length
45269	47480:	contig of 2212 bp in length
47481	47580:	gap of unknown length
47581	48906:	contig of 1326 bp in length
48907	49006:	gap of unknown length
49007	50197:	contig of 1191 bp in length
50198	50297:	gap of unknown length
50299	51599:	contig of 1302 bp in length
51600	51699:	gap of unknown length
51700	52929:	contig of 1230 bp in length
52930	53029:	gap of unknown length
53030	55240:	contig of 2211 bp in length
55241	55340:	gap of unknown length
55341	56767:	contig of 1427 bp in length
56768	56867:	gap of unknown length
56868	58429:	contig of 1562 bp in length
58430	58529:	gap of unknown length
58530	59726:	contig of 1197 bp in length
59727	59826:	gap of unknown length
59827	61354:	contig of 1528 bp in length
61355	61454:	gap of unknown length
61455	63377:	contig of 1883 bp in length
63378	63437:	gap of unknown length
63438	64756:	contig of 1319 bp in length
64757	64856:	gap of unknown length
64857	66353:	contig of 1497 bp in length
66354	66453:	gap of unknown length
66454	67492:	contig of 1039 bp in length
67493	67592:	gap of unknown length
67593	69374:	contig of 1782 bp in length
69375	69474:	gap of unknown length
69475	71849:	contig of 2375 bp in length
71850	71949:	gap of unknown length
71950	73590:	contig of 1641 bp in length
73591	73690:	gap of unknown length
73691	75179:	contig of 1489 bp in length
75180	75279:	gap of unknown length
75280	78068:	contig of 2789 bp in length
78069	78168:	gap of unknown length
78169	80552:	contig of 2384 bp in length
80553	80652:	gap of unknown length
80653	82396:	contig of 1744 bp in length

```

* 82397 82496: gap of unknown length
* 82497 83902: contig of 1406 bp in length

Query Match
Best Local Similarity 16.6%; Score 34.2; DB 2; Length 156460;
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 7 AGCTCGTACCATCTGATGCTCAGAGACCAAGTAAAGTCTGCTGCTACTAAGC 66
DB 5568 AGATATGTTAAATTTTGTATATATTAAGATAATATTCATTTTGTCTTTAACTAATC 5627

QY 67 CAGGTTCTTCTTATATCTTATCTGTTGCTGCTATGTTAAACCACTTAACCGTTGTT 126
DB 5628 TATTTGTATAGTTTATTTACTCTTATGCTTATGCTTATGAACTCCCTGTCATTTCAG 5687

QY 127 TGAAGGACACTGATGCTCAGGATCAAAAAGTCTGCTGAAGTTCC 173
DB 5688 AGATAGAAACAGCATATTCAAAATTCATTAATCTTTTGTATGATATAC 5734

AC100679 61448 bp DNA linear HTG 22-NOV-2001
Mus musculus clone RP23-16701, LOW-PASS SEQUENCE SAMPLING.
AC100679
AC100679.1 GI:17048045
HTG: HTGS, PHASE0.
Mus musculus.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 61448)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-16701
Unpublished
2 (bases 1 to 61448)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campioniano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15938
Center clone name: 167_0_1
-----

```

```

* NOTE: This record contains 76 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 892: contig of 692 bp in length
* 693 792: gap of 100 bp
* 793 1505: contig of 713 bp in length
* 1506 1605: gap of 100 bp
* 1606 2326: contig of 721 bp in length
* 2327 2426: gap of 100 bp
* 2427 3168: contig of 742 bp in length
* 3169 3268: gap of 100 bp
* 3269 3970: contig of 702 bp in length
* 3971 4070: gap of 100 bp
* 4071 4780: contig of 710 bp in length
* 4781 4880: gap of 100 bp
* 4881 5571: contig of 691 bp in length
* 5572 5671: gap of 100 bp
* 5672 6382: contig of 711 bp in length
* 6383 6482: gap of 100 bp
* 6483 7204: contig of 722 bp in length
* 7205 7304: gap of 100 bp
* 7305 8027: contig of 723 bp in length
* 8028 8127: gap of 100 bp
* 8128 8863: contig of 736 bp in length
* 8864 8963: gap of 100 bp
* 8964 9663: contig of 700 bp in length
* 9664 9763: gap of 100 bp
* 9764 10438: contig of 675 bp in length
* 10439 10538: gap of 100 bp
* 10539 11247: contig of 709 bp in length
* 11248 11347: gap of 100 bp
* 11348 12044: contig of 697 bp in length
* 12045 12144: gap of 100 bp
* 12145 12854: contig of 710 bp in length
* 12855 12954: gap of 100 bp
* 12955 13666: contig of 712 bp in length
* 13667 13766: gap of 100 bp
* 13767 14488: contig of 722 bp in length
* 14489 14588: gap of 100 bp
* 14589 15305: contig of 717 bp in length
* 15306 15405: gap of 100 bp
* 15406 16139: contig of 734 bp in length
* 16140 16239: gap of 100 bp
* 16240 16966: contig of 727 bp in length
* 16967 17066: gap of 100 bp
* 17067 17756: contig of 690 bp in length
* 17757 17856: gap of 100 bp
* 17857 18562: contig of 706 bp in length
* 18563 18662: gap of 100 bp
* 18663 19346: contig of 684 bp in length
* 19347 19446: gap of 100 bp
* 19447 20149: contig of 703 bp in length
* 20150 20249: gap of 100 bp
* 20250 20943: contig of 694 bp in length
* 20944 21043: gap of 100 bp
* 21044 21764: contig of 721 bp in length
* 21765 21864: gap of 100 bp
* 21865 22580: contig of 716 bp in length
* 22581 22680: gap of 100 bp
* 22681 23354: contig of 674 bp in length
* 23355 23454: gap of 100 bp
* 23455 24180: contig of 726 bp in length
* 24181 24280: gap of 100 bp
* 24281 25014: contig of 734 bp in length
* 25015 25114: gap of 100 bp

```

* 25115 25815: contig of 701 bp in length
* 25816 25915: gap of 100 bp
* 25916 26595: contig of 680 bp in length
* 26596 26695: gap of 100 bp
* 26696 27403: contig of 708 bp in length
* 27404 27503: gap of 100 bp
* 27504 28224: contig of 721 bp in length
* 28225 28324: gap of 100 bp
* 28325 29056: contig of 732 bp in length
* 29057 29156: gap of 100 bp
* 29157 29885: contig of 729 bp in length
* 29886 29985: gap of 100 bp
* 30669: contig of 684 bp in length
* 30670 30769: gap of 100 bp
* 30770 31456: contig of 687 bp in length
* 31457 31556: gap of 100 bp
* 31557 32285: contig of 729 bp in length
* 32286 32385: gap of 100 bp
* 32386 33116: contig of 731 bp in length
* 33117 33216: gap of 100 bp
* 33217 33910: contig of 694 bp in length
* 33911 34010: gap of 100 bp
* 34011 34709: contig of 699 bp in length
* 34710 34809: gap of 100 bp
* 34810 35487: contig of 678 bp in length
* 35488 35587: gap of 100 bp
* 35588 36291: contig of 704 bp in length
* 36292 36391: gap of 100 bp
* 36392 37074: contig of 683 bp in length
* 37075 37174: gap of 100 bp
* 37175 37875: contig of 701 bp in length
* 37876 37975: gap of 100 bp
* 37976 38773: contig of 798 bp in length
* 38774 38873: gap of 100 bp
* 38874 39591: contig of 718 bp in length
* 39592 39691: gap of 100 bp
* 39692 40406: contig of 715 bp in length
* 40407 40506: gap of 100 bp
* 40507 41188: contig of 682 bp in length
* 41189 41288: gap of 100 bp
* 41289 41941: contig of 653 bp in length
* 41942 42041: gap of 100 bp
* 42042 42711: contig of 670 bp in length
* 42712 42811: gap of 100 bp
* 42812 43523: contig of 712 bp in length
* 43524 43623: gap of 100 bp
* 43624 44343: contig of 720 bp in length
* 44344 44443: gap of 100 bp
* 44444 45156: contig of 713 bp in length
* 45157 45256: gap of 100 bp
* 45257 45948: contig of 692 bp in length
* 45949 46048: gap of 100 bp
* 46049 46772: contig of 724 bp in length
* 46773 46872: gap of 100 bp
* 46873 47598: contig of 726 bp in length
* 47599 47698: gap of 100 bp
* 47699 48417: contig of 719 bp in length
* 48418 48517: gap of 100 bp
* 48518 49217: contig of 700 bp in length
* 49218 49317: gap of 100 bp
* 49318 50003: contig of 686 bp in length
* 50004 50103: gap of 100 bp
* 50104 50816: contig of 713 bp in length
* 50817 50916: gap of 100 bp
* 50917 51631: contig of 715 bp in length
* 51632 51731: gap of 100 bp
* 51732 52444: contig of 713 bp in length
* 52445 52544: gap of 100 bp
* 52545 53260: contig of 716 bp in length
* 53261 53360: gap of 100 bp
* 53361 54078: contig of 718 bp in length
* 54079 54178: gap of 100 bp
* 54179 54861: contig of 683 bp in length

* 54862 54961: gap of 100 bp
* 54962 55700: contig of 739 bp in length
* 55701 55800: gap of 100 bp

Query Match 16.5%; Score 34; DB 2; Length 61448;
Best Local Similarity 56.1%; Pred. No. 7.8;
Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 81 ATTATCTTGATTCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGAT 140
|||||
Db 45697 ATTACATTGATGATTCCTTATTAACCATCCCTGCTATGATGAACCTAT 45638
|||||

QY 141 TGTCCAGGTATCAAAAGTCTGTGAAGTTCCTCGGTATGCTGTTTCGTT 194
|||||
Db 45637 TTTGTCAGGATGAATGATGCTCTGATGTTCTTGGATTCAGTTTGTCAAGTT 45584
|||||

RESULT 57
AL671891/c
LOCUS
DEFINITION
AL671891 254836 bp DNA linear HTG 13-AUG-2002
MUS musculus chromosome X clone RP23-405E21, *** SEQUENCING IN
PROGRESS ***, 9 unordered pieces.
AL671891
VERSION
AL671891.10 GI:22265400
KEYWORDS
HTG; HTGS_PHASE1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
house mouse
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 254836)
Chapman,J.
Direct Submission
Submitted (09-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 15, 2002 this sequence version replaced gi:21998201.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM405E21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 16% of reads
Chemistry: Dye-terminator Big Dye; 83% of reads
Consensus quality: 250716 bases at least Q40
Consensus quality: 252116 bases at least Q30
Consensus quality: 253330 bases at least Q20
Insert size: 254036; sum-of-contigs
Insert size: 193899; 9.8% error; agarose-fp
Quality coverage: 10.38x in Q20 bases; sum-of-contigs Quality
coverage: 15.09x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 63987: contig of 63987 bp in length
* 63988 64087: gap of 100 bp
* 64088 195855: contig of 131768 bp in length
* 195856 195955: gap of 100 bp
* 195956 213844: contig of 17889 bp in length
* 213845 213944: gap of 100 bp
* 213945 217297: contig of 3353 bp in length
* 217298 217397: gap of 100 bp
* 217398 220897: contig of 3500 bp in length
* 220898 220997: gap of 100 bp

REFERENCE

AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-19B12
JOURNAL Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 178652)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galeagan,J., Gardyna,S., Ginde,S., Gordon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kados,B., Kattas,A., Kellis,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,B., Schuback,R., Seaman,S., Schupack,R., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178652)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kattas,A., Kattas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 24, 2002 this sequence version replaced gi:19703221.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22817
Center clone name: 19_B_12
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175241 bases at least Q40
Consensus quality: 176762 bases at least Q30
Consensus quality: 177264 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 177552; sum-of-contigs

Quality coverage: 8.4 in Q20 bases; agarose-fp
Quality coverage: 8.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 816: contig of 816 bp in length
* 817 916: gap of 100 bp
* 917 1985: contig of 1069 bp in length
* 1986 2085: gap of 100 bp
* 2086 2896: contig of 811 bp in length
* 2897 2996: gap of 100 bp
* 2997 7223: contig of 4227 bp in length
* 7224 7323: gap of 100 bp
* 7324 13183: contig of 5860 bp in length
* 13184 13283: gap of 100 bp
* 13284 126261: contig of 112978 bp in length
* 126262 126361: gap of 100 bp
* 126362 133596: contig of 7235 bp in length
* 133597 133614: contig of 5918 bp in length
* 133615 139714: gap of 100 bp
* 139715 145185: contig of 5471 bp in length
* 145186 145285: gap of 100 bp
* 145286 154037: contig of 8752 bp in length
* 154038 154137: gap of 100 bp
* 154138 175237: contig of 21100 bp in length
* 175238 175337: gap of 100 bp
* 175338 178652: contig of 3315 bp in length.

FEATURES

Source

1..178652
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-19B12"
/clone_lib="RPCI-23 Female Mouse BAC"

misc_feature

1..816
/note="assembly_fragment"
clone_end:SP6
vector_side:left

misc_feature

917..1985
/note="assembly_fragment"

misc_feature

2086..2896
/note="assembly_fragment"

misc_feature

2997..7223
/note="assembly_fragment"

misc_feature

7324..13183
/note="assembly_fragment"

misc_feature

13284..126261
/note="assembly_fragment"

misc_feature

126362..133596
/note="assembly_fragment"

misc_feature

133615..139714
/note="assembly_fragment"

misc_feature

139715..145185
/note="assembly_fragment"

misc_feature

145286..154037
/note="assembly_fragment"

misc_feature

154138..175237
/note="assembly_fragment"

misc_feature

175338..178652
/note="assembly_fragment"

misc_feature

clone_end:T7
vector_side:right

BASE COUNT 53824 a 33935 c 32001 g 57792 t 1100 others
ORIGIN

Query Match

16.4% Score 33.8; DB 2; Length 178652;

Best Local Similarity 55.6%; Pred. No. 8.8;
Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAAACGTTGTTGAAGGACACTGATTGTCACG 147
|||||
Db 102074 TGATTCCTCTGTTACTATCAACTGTGAACCTTAAACATCTGACAGTGAACCAAG 102015
|||||
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGGGTATGCTTTGTTTCGTTCCACAATAAT 204
|||||
Db 102014 CTTTCTCCATCTCTGGAAGGCTTTGAGGCAAGGTTATGATGTTCCACTATACT 101958
|||||

RESULT 60
AF002719 2435 bp DNA linear ROD 26-JAN-1999
LOCUS Mus musculus secretory leukoprotease inhibitor gene, complete cds.
DEFINITION AF002719
ACCESSION AF002719
VERSION AF002719.1 GI:4100898
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2435)
AUTHORS Kikuchi,T., Abe,T., Hoshi,S., Matsubara,N., Tominaga,Y., Satoh,K.
and Nukiwa,T.
TITLE Structure of the murine secretory leukoprotease inhibitor (Slpi)
gene and chromosomal localization of the human and murine SLPI
genes
JOURNAL Am. J. Respir. Cell Mol. Biol. 19 (6), 875-880 (1998)
MEDLINE 99061820
PUBMED 9843921
REFERENCE 2 (bases 1 to 2435)
AUTHORS Kikuchi,T., Abe,T., Hoshi,S., Matsubara,N., Tominaga,Y., Satoh,K.
and Nukiwa,T.
TITLE Direct Submission
Submitted (06-MAY-1997) Department of Respiratory Oncology and
Molecular Medicine, Institute of Development, Aging and Cancer,
Tohoku University, 4-1 Seiryomachi, Aoba-ku, Sendai, Miyagi-ken
980-77, Japan

FEATURES
source
1. .2435
/organism="Mus musculus"
/strain="129SV"
/db_xref="taxon:10090"
/chromosome="2"
join(101..205,781..942,1368..1517,2080..2322)
/product="secretory leukoprotease inhibitor"
join(121..205,781..942,1368..1516)
/codon_start=1
/product="secretory leukoprotease inhibitor"
/protein_id="AAD09307.1"
/db_xref="GI:4100898"
/translation="MKSGLLPFTVLLALGILAPVTVEGKNDAIKIGACPAKPAQC
LKLKPEQRTDMECPQKRCQDAGCKVPIRKPVWRKPCVKTQARCMMLNP
PNCQDQGDGKCKCEGICGKVLPPM"
BASE COUNT 509 a 600 c 652 g 674 t

Query Match 16.3%; Score 33.6; DB 10; Length 2435;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 52 CTGTGCTACTAAGCCAGGTTCTTGCTCTATTATCTGTTGCGCTATGTTAAACC 111
|||||
Db 1365 CAGTGGAGGAGCCCTGGGAGGTGCGTCAAACTCAGGCAAGATGATGCTTAACC 1424
|||||
QY 112 CACCTAACCGTTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGTGCAAGGTT 171
|||||
Db 1425 CTCCCATGTCGCAGGAGGACGGGAGTGTGACGCGCAATACAGTGTGTGAGGGA 1484
|||||
QY 172 CCTCGGTATGGCTTG 187
|||||

Db 1485 TATGTGGGAAAGTCTG 1500

RESULT 61
AL662897 84987 bp DNA linear PRI 22-JAN-2002
LOCUS Human DNA sequence from clone RP11-309C8 on chromosome X, complete
DEFINITION sequence.
ACCESSION AL662897
VERSION AL662897.6 GI:18307360
KEYWORDS HTG.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 84987)
AUTHORS Howden,P.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 23, 2002 this sequence version replaced gi:18250572.

COMMENT During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep

This sequence was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX

RP11-309C8 is from the library RPc1-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone
RP11-309C8 it may be shorter because we sequence overlapping
sections only once, except for a short overlap.

The true left end of clone RP1-198G23 is at 82988 in this sequence.
The true right end of clone RP11-378H3 is at 2000 in this sequence.

FEATURES
source
1. .84987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-309C8"
/clone_lib="RPc1-11.2"
BASE COUNT 26631 a 15105 c 15585 g 27666 t

ORIGIN
Query Match 16.3%; Score 33.6; DB 9; Length 84987;
Best Local Similarity 57.7%; Pred. No. 10;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 53 TGTCGTCTACTAAGCCAGGTTCTTGCTCTATTATCTGTTGCGCTATGTTAAACC 112
|||||
Db 49793 TCTGCTTTTATTCTTATTATGCGGTATATTATGTTGATTGTTTTCATGTTAAACCA 49852
|||||
QY 113 ACCTAACCGTTGTTTGAAGGACACTGATTGTCAGGTATCAAAA 156
|||||


```

* 7587 10465: contig of 2879 bp in length
* 10466 10565: gap of unknown length
* 10566 12280: contig of 1715 bp in length
* 12281 12380: gap of unknown length
* 12381 15032: contig of 2652 bp in length
* 15033 15132: gap of unknown length
* 15133 17864: contig of 2732 bp in length
* 17865 17965: gap of unknown length
* 17965 23051: contig of 5087 bp in length
* 23052 23151: gap of unknown length
* 23152 27061: contig of 3910 bp in length
* 27062 27161: gap of unknown length
* 27162 32335: contig of 5174 bp in length
* 32336 32435: gap of unknown length
* 32436 37651: contig of 5216 bp in length
* 37652 37751: gap of unknown length
* 37752 42402: contig of 4651 bp in length
* 42403 42502: gap of unknown length
* 42503 46527: contig of 4025 bp in length
* 46528 46627: gap of unknown length
* 46628 52207: contig of 5580 bp in length
* 52208 52307: gap of unknown length
* 52308 58003: contig of 5696 bp in length
* 58004 58104: gap of unknown length
* 58104 67234: contig of 9131 bp in length
* 67235 76254: gap of unknown length
* 76255 76354: gap of unknown length
* 76355 89495: contig of 13140 bp in length
* 89496 89594: gap of unknown length
* 89595 104109: contig of 14515 bp in length
* 104110 104209: gap of unknown length
* 104210 118259: contig of 14050 bp in length
* 118260 118359: gap of unknown length
* 118360 138555: contig of 20196 bp in length
* 138556 138656: gap of unknown length
* 138657 160851: contig of 22196 bp in length
* 160852 160952: gap of unknown length
* 160953 193302: contig of 32351 bp in length.

```

FEATURES

Source

```

1. 193302
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-561H2"
1. .1164
/note="assembly_name:Contig21"
1265. .2782
/note="assembly_name:Contig24"
2883. .4288
/note="assembly_name:Contig25"
4389. .6059
/note="assembly_name:Contig26"
6160. .7486
/note="assembly_name:Contig27"
7587. .10465
/note="assembly_name:Contig28"
10566. .12280
/note="assembly_name:Contig29"
12381. .15032
/note="assembly_name:Contig30"
15133. .17864
/note="assembly_name:Contig31"
17965. .23051
/note="assembly_name:Contig32"
23152. .27061
/note="assembly_name:Contig33"
27162. .32335
/note="assembly_name:Contig34"
32436. .37651
/note="assembly_name:Contig35"
37752. .42402
/note="assembly_name:Contig36"

```

```

misc_feature 42503. .46527
/note="assembly_name:Contig37"
misc_feature 46628. .52207
/note="assembly_name:Contig38"
misc_feature 52308. .58003
/note="assembly_name:Contig39"
misc_feature 58104. .67234
/note="assembly_name:Contig40"
misc_feature 67335. .76254
/note="assembly_name:Contig41"
misc_feature 76355. .89494
/note="assembly_name:Contig42"
misc_feature 89595. .104109
/note="assembly_name:Contig43"
misc_feature 104210. .118259
/note="assembly_name:Contig44"
misc_feature 118360. .138555
/note="assembly_name:Contig45
clone_end:SP6
vector_side:left"
misc_feature 138656. .160851
/note="assembly_name:Contig46"
misc_feature 160952. .193302
/note="assembly_name:Contig47"
BASE COUNT 61908 a 35369 c 35618 g 57952 t 2455 others
ORIGIN

```

```

Query Match 16.3%; Score 33.6; DB 2; Length 193302;
Best Local Similarity 57.7%; Pred. No. 10;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

Qy 53 TGTCTCTACTAAGCCAGGTCCTCTCTATATCTGATTCGTCGCGCTATGTTAAACCC 112
Db 131012 TCTGCTTTTATCTATTTATGCGGTATATATGTTGATTTGTCATGTTAAACCA 130953

Qy 113 ACCTAACCGTGTGTTGAAGACACTGATTCCTCAGGATCAAAA 156
Db 130952 AACTTACATTTTGGGATAAATAATAATGTCATTTTACATTA 130909

```

RESULT 64

HS319F24

```

LOCUS HS319F24 28515 bp DNA linear PRI 12-DEC-1999
DEFINITION Human DNA sequence from clone RP1-319F24 on chromosome
22q12.3-13.31 Contains a GSS, complete sequence.

```

ACCESSION

AL035495

VERSION

AL035495.13 GI:5050958

KEYWORDS

HTG.

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 28515)

TITLE

Sehra,H

JOURNAL

Direct Submission

Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerquest@sanger.ac.uk

On Jun 11, 1999 this sequence version replaced gi:5002615.

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

This sequence was generated from part of bacterial clone contigs of

human chromosome 22, constructed by the Sanger Centre Chromosome 22

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr22

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP1-319F24 is from the library RPI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pCVPAC2

IMPORTANT: This sequence is not the entire insert of clone RP1-319F24. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone RP3-508I15 is at 28416 in this sequence.
 The true right end of clone RP1-199H16 is at 100 in this sequence.

FEATURES

source

```

1..28515
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="22"
  /map="q12.3-13.31"
  /clone="RP1-319F24"
  /clone_lib="RPC1-1"
2..157
  /note="AluX repeat: matches 1..154 of consensus"
  1186..1231
    /note="L2 repeat: matches 1747..2746 of consensus"
  1233..1534
    /note="L2 repeat: matches 1747..2746 of consensus"
  complement(1558..1698)
    /note="Single clone region"
  1573..1867
    /note="AluX repeat: matches 1..302 of consensus"
  1926..2825
    /note="L2 repeat: matches 1747..2746 of consensus"
  3082..3285
    /note="L2 repeat: matches 1747..2746 of consensus"
  3763..4070
    /note="L2 repeat: matches 1747..2746 of consensus"
  4108..4401
    /note="L2 repeat: matches 1747..2746 of consensus"
  4944..5249
    /note="L2 repeat: matches 1747..2746 of consensus"
  5463..5542
    /note="L2 repeat: matches 1747..2746 of consensus"
  5543..5849
    /note="L2 repeat: matches 1747..2746 of consensus"
  5890..6084
    /note="L2 repeat: matches 1747..2746 of consensus"
  6281..6486
    /note="L2 repeat: matches 1747..2746 of consensus"
  6487..6551
    /note="L2 repeat: matches 1747..2746 of consensus"
  6547..6610
    /note="L2 repeat: matches 1747..2746 of consensus"
  6863..6908
    /note="L2 repeat: matches 1747..2746 of consensus"
  6952..7069
    /note="L2 repeat: matches 1747..2746 of consensus"
  7070..7342
    /note="L2 repeat: matches 1747..2746 of consensus"
  7343..7499
    /note="L2 repeat: matches 1747..2746 of consensus"
  7794..8096
    /note="L2 repeat: matches 1747..2746 of consensus"
  8386..8421
    /note="L2 repeat: matches 1747..2746 of consensus"
  8570..8684
    /note="L2 repeat: matches 1747..2746 of consensus"
  8814..8912
    /note="L2 repeat: matches 1747..2746 of consensus"
  9233..9250
    /note="L2 repeat: matches 1747..2746 of consensus"

```

```

repeat_region
  8944..9016
    /note="L2 repeat: matches 1747..2746 of consensus"
  9017..9241
    /note="L2 repeat: matches 1747..2746 of consensus"
  9245..9331
    /note="L2 repeat: matches 1747..2746 of consensus"
  9332..9625
    /note="L2 repeat: matches 1747..2746 of consensus"
  9626..9683
    /note="L2 repeat: matches 1747..2746 of consensus"
  9684..9989
    /note="L2 repeat: matches 1747..2746 of consensus"
  9990..10296
    /note="L2 repeat: matches 1747..2746 of consensus"
  10305..10441
    /note="L2 repeat: matches 1747..2746 of consensus"
  10656..10969
    /note="L2 repeat: matches 1747..2746 of consensus"
  10971..11220
    /note="L2 repeat: matches 1747..2746 of consensus"
  11393..11517
    /note="L2 repeat: matches 1747..2746 of consensus"
  11525..11617
    /note="L2 repeat: matches 1747..2746 of consensus"
  11678..11974
    /note="L2 repeat: matches 1747..2746 of consensus"
  12666..13415
    /note="L2 repeat: matches 1747..2746 of consensus"
  complement(14319..14778)
    /note="match: GSS: Em:AQ164414"
  14607..14767
    /note="L2 repeat: matches 1747..2746 of consensus"
  14776..14976
    /note="L2 repeat: matches 1747..2746 of consensus"
  14779..15152
    /note="L2 repeat: matches 1747..2746 of consensus"
  14978..15035
    /note="L2 repeat: matches 1747..2746 of consensus"
  15807..15876
    /note="L2 repeat: matches 1747..2746 of consensus"
  15870..15956
    /note="L2 repeat: matches 1747..2746 of consensus"
  15957..16218
    /note="L2 repeat: matches 1747..2746 of consensus"
  16219..16794
    /note="L2 repeat: matches 1747..2746 of consensus"
  16812..17288
    /note="L2 repeat: matches 1747..2746 of consensus"
  17289..17584
    /note="L2 repeat: matches 1747..2746 of consensus"
  17585..18113
    /note="L2 repeat: matches 1747..2746 of consensus"
  18076..18162
    /note="L2 repeat: matches 1747..2746 of consensus"
  18190..18491
    /note="L2 repeat: matches 1747..2746 of consensus"
  18616..18869
    /note="L2 repeat: matches 1747..2746 of consensus"
  19009..19579
    /note="L2 repeat: matches 1747..2746 of consensus"
  19628..19839
    /note="L2 repeat: matches 1747..2746 of consensus"
  19981..20210
    /note="L2 repeat: matches 1747..2746 of consensus"
  20211..20520
    /note="L2 repeat: matches 1747..2746 of consensus"
  20521..20795
    /note="L2 repeat: matches 1747..2746 of consensus"
  21420..21726
    /note="L2 repeat: matches 1747..2746 of consensus"
  21952..22274
    /note="L2 repeat: matches 1747..2746 of consensus"
  22393..22509
    /note="L2 repeat: matches 1747..2746 of consensus"

```

```
repeat_region /note="AluSg/x repeat: matches 174. .295 of consensus"
22613. .22713
/note="MERSA repeat: matches 90. .189 of consensus"
complement(22812. .23205)
/note="match: GSS: Em:AQ278967"
22883. .23289
/note="MER66A repeat: matches 1. .478 of consensus"
23291. .23377
/note="L2 repeat: matches 2612. .2705 of consensus"
23388. .23490
/note="MIR repeat: matches 86. .201 of consensus"
23491. .23796
/note="AluX repeat: matches 1. .308 of consensus"
23797. .23845
/note="MIR repeat: matches 201. .246 of consensus"
24205. .24336
/note="FLAM_C repeat: matches 1. .133 of consensus"
24398. .24504
/note="MIR repeat: matches 48. .152 of consensus"
24594. .24727
/note="L2 repeat: matches 2577. .2735 of consensus"
25034. .25299
/note="AluSg repeat: matches 15. .296 of consensus"

Query Match 16.2%; Score 33.4; DB 9; Length 28515;
Best Local Similarity 55.7%; Pred. No. 12;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 74 TTCTCTATATCTGATTCGTCCTATGTTAAACCCACCTACCGTTGTTGAAGCA 133
|||||
Db 14554 TTTTCTATCTCCAGTCGTGTTCTCTAGTACCACCTTGTTATATGATTGAAGC 14613
|||||

QY 134 CACTGATGTCAGGATCAAAAAGTGTGAAGTCTCTCGCGTATGCGTGTGT 188
|||||
Db 14614 TAATGTTATGAGTCTTACTCTGCTATGCTGTTGCTGAGTCTGTCATGT 14668
|||||

RESULT 65
AC104349
LOCUS Homo sapiens chromosome 8 clone RP11-1058E13 map 8, LOW-PASS
DEFINITION Homo sapiens chromosome 8 clone RP11-1058E13 map 8, LOW-PASS
AC104349
VERSION AC104349.1 GI:17426338
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 64360)
Birken, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 8, clone RP11-1058E13
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 64360)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gatzdyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
```

```
Topham, K., Travers, M., Travis, N., Travis, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21826
Center clone name: 1058_E_13
-----
* NOTE: This record contains 82 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 683: contig of 683 bp in length
* 784 783: gap of 100 bp
* 1435 1434: contig of 651 bp in length
* 1535 1534: gap of 100 bp
* 2232 2231: contig of 697 bp in length
* 2332 2331: gap of 100 bp
* 2999 2999: contig of 668 bp in length
* 3000 3099: gap of 100 bp
* 3100 3785: contig of 686 bp in length
* 3786 3885: gap of 100 bp
* 3886 4553: contig of 668 bp in length
* 4554 4653: gap of 100 bp
* 4654 5346: contig of 693 bp in length
* 5347 5446: gap of 100 bp
* 5447 6129: contig of 683 bp in length
* 6130 6229: gap of 100 bp
* 6230 6919: contig of 690 bp in length
* 6920 7019: gap of 100 bp
* 7020 7700: contig of 681 bp in length
* 7701 7800: gap of 100 bp
* 7801 8488: contig of 688 bp in length
* 8489 8588: gap of 100 bp
* 8589 9261: contig of 673 bp in length
* 9262 9361: gap of 100 bp
* 9362 10027: contig of 666 bp in length
* 10028 10127: gap of 100 bp
* 10128 10806: contig of 679 bp in length
* 10807 10906: gap of 100 bp
* 10907 11574: contig of 668 bp in length
* 11575 11674: gap of 100 bp
* 11675 12366: contig of 692 bp in length
* 12367 12466: gap of 100 bp
* 12467 13160: contig of 694 bp in length
* 13161 13260: gap of 100 bp
* 13261 13966: contig of 706 bp in length
* 13967 14066: gap of 100 bp
* 14067 14756: contig of 690 bp in length
* 14757 14856: gap of 100 bp
* 14857 15557: contig of 701 bp in length
* 15558 15657: gap of 100 bp
* 15658 16339: contig of 682 bp in length
* 16340 16439: gap of 100 bp
* 16440 17100: contig of 661 bp in length
* 17101 17200: gap of 100 bp
```

TITLE
JOURNAL
COMMENT

* 17201 17898: contig of 698 bp in length
 * 17899 17998: gap of 100 bp
 * 17999 18679: contig of 681 bp in length
 * 18680 18779: gap of 100 bp
 * 18780 19451: contig of 672 bp in length
 * 19452 19551: gap of 100 bp
 * 19552 20253: contig of 702 bp in length
 * 20254 20353: gap of 100 bp
 * 20354 21010: contig of 657 bp in length
 * 21011 21110: gap of 100 bp
 * 21111 21802: contig of 692 bp in length
 * 21803 21902: gap of 100 bp
 * 21903 22604: contig of 702 bp in length
 * 22605 22704: gap of 100 bp
 * 22705 23396: contig of 692 bp in length
 * 23397 23496: gap of 100 bp
 * 23497 24187: contig of 691 bp in length
 * 24188 24287: gap of 100 bp
 * 24288 24984: contig of 697 bp in length
 * 24985 25084: gap of 100 bp
 * 25085 25771: contig of 687 bp in length
 * 25772 25871: gap of 100 bp
 * 25872 26542: contig of 671 bp in length
 * 26543 26642: gap of 100 bp
 * 26643 27341: contig of 699 bp in length
 * 27342 27441: gap of 100 bp
 * 27442 28133: contig of 692 bp in length
 * 28134 28233: gap of 100 bp
 * 28234 28909: contig of 676 bp in length
 * 28910 29009: gap of 100 bp
 * 29010 29708: contig of 699 bp in length
 * 29709 29808: gap of 100 bp
 * 29809 30491: contig of 683 bp in length
 * 30492 30591: gap of 100 bp
 * 30592 31273: contig of 682 bp in length
 * 31274 31373: gap of 100 bp
 * 31374 32067: contig of 694 bp in length
 * 32068 32167: gap of 100 bp
 * 32168 32858: contig of 691 bp in length
 * 32859 32958: gap of 100 bp
 * 32959 33624: contig of 666 bp in length
 * 33625 33724: gap of 100 bp
 * 33725 34427: contig of 703 bp in length
 * 34428 34527: gap of 100 bp
 * 34528 35195: contig of 668 bp in length
 * 35196 35295: gap of 100 bp
 * 35296 35968: contig of 673 bp in length
 * 35969 36068: gap of 100 bp
 * 36069 36751: contig of 683 bp in length
 * 36752 36851: gap of 100 bp
 * 36852 37532: contig of 681 bp in length
 * 37533 37632: gap of 100 bp
 * 37633 38361: contig of 729 bp in length
 * 38362 38461: gap of 100 bp
 * 38462 39151: contig of 690 bp in length
 * 39152 39251: gap of 100 bp
 * 39252 39942: contig of 691 bp in length
 * 39943 40042: gap of 100 bp
 * 40043 40731: contig of 689 bp in length
 * 40732 40831: gap of 100 bp
 * 40832 41519: contig of 688 bp in length
 * 41520 41619: gap of 100 bp
 * 41620 42297: contig of 678 bp in length
 * 42298 42397: gap of 100 bp
 * 42398 43082: contig of 685 bp in length
 * 43083 43182: gap of 100 bp
 * 43183 43870: contig of 688 bp in length
 * 43871 43970: gap of 100 bp
 * 43971 44638: contig of 668 bp in length
 * 44639 44738: gap of 100 bp
 * 44739 45422: contig of 684 bp in length
 * 45423 45522: gap of 100 bp
 * 45523 46220: contig of 698 bp in length

* 46221 46320: gap of 100 bp
 * 46321 46985: contig of 665 bp in length
 * 46986 47085: gap of 100 bp
 * 47086 47761: contig of 676 bp in length
 * 47762 47861: gap of 100 bp
 * 47862 48548: contig of 687 bp in length
 * 48549 48648: gap of 100 bp
 * 48649 49341: contig of 693 bp in length
 * 49342 49441: gap of 100 bp
 * 49442 50130: contig of 689 bp in length
 * 50131 50230: gap of 100 bp
 * 50231 50906: contig of 676 bp in length
 * 50907 51006: gap of 100 bp
 * 51007 51703: contig of 697 bp in length
 * 51704 51803: gap of 100 bp
 * 51804 52494: contig of 691 bp in length
 * 52495 52594: gap of 100 bp
 * 52595 53288: contig of 694 bp in length
 * 53289 53388: gap of 100 bp
 * 53389 54077: contig of 689 bp in length

Query Match 16.2%; Score 33.4; DB 2; Length 64360;

Best Local Similarity 55.7%; Pred. No. 12;

Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 74 TTGCTCTATTATCTGATTCTGTTGCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGA 133
 |||||
 Db 56071 TTTTCCTAATCTCCAGCTGTGTTCTTAGGTACCCACCTGTTATATAATGATTAAAGAC 56130

QY 134 CACTGATTCTCCAGGTATCAAAAAGTGTGCTGAAGTCTCTCGGTATGCGCTTGT 188
 |||||
 Db 56131 TTAATGTTTATTGAGTCTTACTCTGTGCTATGCTGTTTGTGAGTGTGCTTGCATGT 56185

RESULT 66

AC107719/c

LOCUS AC107719 232793 bp DNA linear HTG 05-JUN-2002
 DEFINITION Mus musculus clone RP23-187P2, WORKING DRAFT SEQUENCE, 12 ordered pieces.

AC107719

VERSION AC107719.2 GI:21326309

HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS house mouse.

SOURCE

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

1 (bases 1 to 232793)

Mus musculus, clone RP23-187P2

Unpublished

2 (bases 1 to 232793)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,I., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,W., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNALREFERENCE
AUTHORS

Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 232793)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J. J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneses, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 5, 2002 this sequence version replaced gi:18308421.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

TITLE
JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19588

Center clone name: 187_P2

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 229008 bases at least Q40

Consensus quality: 230799 bases at least Q30

Consensus quality: 231389 bases at least Q20

Insert size: 240000; agarose-fp

Insert size: 231693; sum-of-contigs

Quality coverage: 8.0 in Q20 bases; agarose-fp

Quality coverage: 8.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced

by the finished sequence as soon as it is available and the accession number will be preserved.

1 31998: contig of 31998 bp in length

31999 32098: gap of 100 bp

32099 33165: contig of 1067 bp in length

33166 33265: gap of 100 bp

33266 35741: contig of 2476 bp in length

35742 35841: gap of 100 bp

35842 38517: contig of 2676 bp in length

38518 38617: gap of 100 bp

38618 46169: contig of 7552 bp in length

46170 46269: gap of 100 bp

46270 101040: contig of 54771 bp in length

* 101041 101140: gap of 100 bp
* 101141 116926: contig of 15786 bp in length
* 116927 117026: gap of 100 bp
* 117027 135879: contig of 18853 bp in length
* 135880 135979: gap of 100 bp
* 135980 164306: contig of 28327 bp in length
* 164307 164406: gap of 100 bp
* 164407 194369: contig of 29963 bp in length
* 194370 194469: gap of 100 bp
* 194470 223647: contig of 29178 bp in length
* 223648 223747: gap of 100 bp
* 223748 232793: contig of 9046 bp in length.

FEATURES

Location/Qualifiers

source
1..232793
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-187P2"
/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature
1..31998
/note="assembly_fragment"
clone_end:SP6
vector_side:left
32099..33165
/note="assembly_fragment"
misc_feature
33266..35741
/note="assembly_fragment"
misc_feature
35842..38517
/note="assembly_fragment"
misc_feature
38618..46169
/note="assembly_fragment"
misc_feature
46270..101040
/note="assembly_fragment"
misc_feature
101141..116926
/note="assembly_fragment"
misc_feature
117027..135879
/note="assembly_fragment"
misc_feature
135980..164306
/note="assembly_fragment"
misc_feature
164407..194369
/note="assembly_fragment"
misc_feature
194470..223647
/note="assembly_fragment"
misc_feature
223748..232793
/note="assembly_fragment"
clone_end:T7
vector_side:right

BASE COUNT 72773 a 45524 c 45372 g 68024 t 1100 others
ORIGIN

Query Match 16.2%; Score 33.4; DB 2; Length 232793;
Best Local Similarity 57.0%; Pred. No. 12;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 81 ATTATCTGATTCCTGGCTATGCTTAACCCACCTAACCGTTGTTTGAAGGACACTGAT 140
Db 167215 AATATGTTGATGGATTTCCTTATATTAACCATCCCTGGATCCCTGGATGAACCTACT 167216
Qy 141 TGTCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGGTATGGCTTG 187
Db 167215 TGATCAGGATTAACGATGGCTTTGAAGTGTCTTGGATTGGTTG 167169

RESULT 67

AF205374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mus musculus.

Mus musculus.

Mus musculus.

Mus musculus.

Mus musculus.

Mus musculus.

Mus musculus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 4090)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Lei.K.-J., Jin,W.-W. and Wahl,S.M.
Isolation and characterization of the gene for secretory leukocyte
protease inhibitor from 129/SvJ mice
Unpublished

2 (bases 1 to 4090)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Lei.K.-J., Jin,W.-W. and Wahl,S.M.
Direct Submission
Submitted (16-NOV-1999) Oral Infection and Immunity Branch,
National Institute of Dental & Craniofacial Research, 30 Convent
Dr., MSC4352, Bldg. 30, Rm. B06, Bethesda, MD 20892, USA
Location/Qualifiers
1. .4090
/organism="Mus musculus"
/strain="129/SvJ"
/db_xref="taxon:10090"
1266..2710
/gene="Slpi"
Join(1266..1353,1923..2080,2561..2710)
/gene="Slpi"
/product="secretory leukocyte protease inhibitor"
Join(1266..1353,1923..2080,2561..2710)
/gene="Slpi"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAF19194.1"
/db_xref="GI:6606529"
/translation="MKSGLLPFTVLLALGILAPWTVEGKNDKIKGACPAKKPAQC
LKLEKPCRTWECPCGKRCOCADGSKCNPVIRKPVWRKPGRCVKYQARCMNLN
PNVQRDQCQDGKXKCEGICGKVCLEPDM"
BASE COUNT 944 a 953 c 1069 g 1118 t 6 others
ORIGIN

Query Match 16.1%; Score 33.2; DB 10; Length 4090;
Best Local Similarity 53.0%; Pred. No. 14;
Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 54 GTGCTACTAAGCAGGTCTTCTATATCTTGTGTCGCTATGTTAAACCCA 113
DB 2561 GTGTGGAGGAGCCTGGGAGTGGCGTCAAAACTCAGGCAAGATGATGATGTTAACCT 2620
QY 114 CTAACCGTTTGAAGGACACTGATGTCAGGTATCAAAAGTGTGTAAGGTTC 173
DB 2621 CCAATGTCGCAGGAGCGGCGAGTGTGACGCAATAACAAGTGTGTGAGGGTATA 2680
QY 174 TGGGTATGGCTTG 187
DB 2681 TGTGGAAAGTCTG 2694

RESULT 68
AC126796
LOCUS
DEFINITION
AC126796 236121 bp DNA linear HTG 10-JUL-2002
SEQUENCE, 6 unordered pieces.
AC126796
VERSION
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 236121)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished

2 (bases 1 to 236121)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (09-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
Location/Qualifiers
1. .236121
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-384F19"

BASE COUNT 71322 a 46827 c 47077 g 70356 t 539 others
ORIGIN

Query Match 16.1%; Score 33.2; DB 2; Length 236121;
Best Local Similarity 59.6%; Pred. No. 13;
Matches 56; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 27 TATGCTCAAGAACCACTTAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTATTATC 86

McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (10-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Jul 10, 2002 this sequence version replaced gi:21717262.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M.BA0384F19
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 233502 bases at least Q40
Consensus quality: 233911 bases at least Q30
Consensus quality: 234236 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 235621; sum-of-contigs
Quality coverage: 0.00 in Q20 bases; agarose-fp
Quality coverage: 11.87 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 7176: contig of 7176 bp in length
* 7177 7276: gap of unknown length
* 7277 18172: contig of 10896 bp in length
* 18173 18272: gap of unknown length
* 18273 33690: contig of 15418 bp in length
* 33691 33790: gap of unknown length
* 33791 94883: contig of 61093 bp in length
* 94884 160810: contig of 65827 bp in length
* 160811 160910: gap of unknown length
* 160911 236121: contig of 75211 bp in length.
* Location/Qualifiers
1. .236121
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-384F19"
misc_feature 1. .7176
/note="assembly_name:Contig2"
misc_feature 7277..18172
/note="assembly_name:Contig3"
misc_feature 18273..33690
/note="assembly_name:Contig4"
misc_feature 33791..94883
/note="assembly_name:Contig5"
misc_feature 94884..160810
/note="assembly_name:Contig6"
misc_feature 160911..236121
/note="assembly_name:Contig7"
BASE COUNT 71322 a 46827 c 47077 g 70356 t 539 others
ORIGIN

Query Match 16.1%; Score 33.2; DB 2; Length 236121;
Best Local Similarity 59.6%; Pred. No. 13;
Matches 56; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 27 TATGCTCAAGAACCACTTAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTATTATC 86

```

Db 81022 TTTTCACTAGACTCAAGTCCAGCTCAATCTCTTCCACCAGAGTCTGGCAATTCTC 81081
QY      87 TTTGATTTCGTCGCTATGTTAAACCCACCTAAC 120
Db 81082 TTTGATTTCATCTCTCTGACAAAACAAACAAAC 81115

RESULT 69
AF375596S2      17515 bp DNA linear ROD 20-MAY-2001
LOCUS Mus musculus medium and short chain L-3-hydroxyacyl-Coenzyme A
DEFINITION dehydrogenase (Mschad) gene, exons 2 through 8, and complete cds;
nuclear gene for mitochondrial product.
ACCESSION AF375597
VERSION AF375597.1 GI:14150816
SEGMENT 2 of 2
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 17515)
AUTHORS O'Brien,L.K., Sims,H.F. and Strauss,A.W.
TITLE Mouse medium and short chain L-3-hydroxyacyl-Coenzyme A
dehydrogenase gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 17515)
AUTHORS O'Brien,L.K., Sims,H.F. and Strauss,A.W.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) Pediatrics, Vanderbilt Medical Center, 1161
21st Avenue South, D-2220 MCN, Nashville, TN 37232, USA
FEATURES
source 1. .17515
/organism="Mus musculus"
/db_xref="taxon:10090"
/gene="Mschad"
join(AF375596:1:6458..10094,1..17515)
5986..6112,8631..8720,10231..10303,15582..15698,
16954..17515)
/genes="Mschad"
/product="medium and short chain L-3-hydroxyacyl-Coenzyme
A dehydrogenase"
CDS join(AF375596:1:6478..6609,1329..1457,2698..2855,
5986..6112,8631..8720,10231..10303,15582..15698,
16954..17072)
/genes="Mschad"
/note="L-3-hydroxyacyl-Coenzyme A dehydrogenase;
NAD-dependent dehydrogenase; specific for medium and short
chain fatty acids; SCHAD; HADH"
/codon_start=1
/product="medium and short chain L-3-hydroxyacyl-Coenzyme
A dehydrogenase"
/protein_id="AAK54642.1"
/db_xref="GI:14150817"
/translation="MAFVTRQFLRSMSSSSAAAKILIKHVTVIGGLMGAGIAQ
VAAATGHTVLVDOTEDILAKSKGIEESLKRMAKKFTENPKADGEVEXTLSCLST
STDAASVHSTDLVVEAIVENLKNELQRLDKFAEHTIFASNTSLQITNAT
TQDRPAGLHFNFPVPMKLVIEVITKPTMSQKTFESLVDKFKLGHVPSCKDTPGFI
VNRLLVPYLIIEAVRLHREGDASKEDIDTAMKLGAGYPMGFELLDYVGLDITTKFILDG
WHMEPEPNLFPQSPSMNVLVAQKLGKKTGSGFYKYK"
1329..1427
/genes="Mschad"
/number=2
2698..2855
/genes="Mschad"
/number=3
5986..6112
/genes="Mschad"
/number=4
8631..8720
/genes="Mschad"

```

```

/number=5
10231..10303
/genes="Mschad"
/number=6
15582..15698
/genes="Mschad"
/number=7
16954..17515
/genes="Mschad"
/number=8

BASE COUNT 3910 a 4109 c 4462 g 5026 t 8 others
ORIGIN

Query Match 16.0%; Score 33; DB 10; Length 17515;
Best Local Similarity 52.6%; Pred. No. 16;
Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 69 GGTTCTTGTCTATTATCTTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTG 128
Db 9488 GTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 9547
QY 129 AAGGACACTGATTCGCCAGGTATCAAAAAGTGTGTAAGGTTCCCTCGGTATGGCTTGT 188
Db 9548 TCCTGAACCATGAGTCAGTCTCTTTAGTACTGGGCCCTCTACTGCCTAATGGCTTGT 9607
QY 189 TTCGTTCCACAATAATA 205
Db 9608 TTCAGAGAACAACAACTA 9624

RESULT 70
AC119403      40223 bp DNA linear PRI 28-JUN-2002
LOCUS Homo sapiens chromosome 19 clone LLNLR-265A7, complete sequence.
DEFINITION Homo sapiens chromosome 19 clone LLNLR-265A7, complete sequence.
ACCESSION AC119403
VERSION AC119403.2 GI:21622702
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40223)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 40223)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 40223)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On June 28, 2002 this sequence version replaced gi:20330817.
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
source 1. .40223
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="LLNLR-265A7"

BASE COUNT 10097 a 9482 c 9319 g 11325 t
ORIGIN

Query Match 16.0%; Score 33; DB 9; Length 40223;
Best Local Similarity 54.5%; Pred. No. 16;

```


COMMENT

On May 4, 2002 this sequence version replaced gi:20389394.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center Project name: L19584

Center Clone name: 157.H.14

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 200834 bases at least Q40

Consensus quality: 204678 bases at least Q30

Consensus quality: 206239 bases at least Q20

Insert size: 188000; agarose-fp

Insert size: 209203; sum-of-contigs

Quality coverage: 8.3 in Q20 bases; agarose-fp

Quality coverage: 7.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 46 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 651: contig of 651 bp in length

* 652 751: gap of 100 bp

* 752 1458: contig of 707 bp in length

* 1459 1558: gap of 100 bp

* 1559 2267: contig of 709 bp in length

* 2268 2367: gap of 100 bp

* 2368 3028: contig of 661 bp in length

* 3029 3128: gap of 100 bp

* 3129 3802: contig of 674 bp in length

* 3803 3902: gap of 100 bp

* 3903 4563: contig of 661 bp in length

* 4564 4663: gap of 100 bp

* 4664 5327: contig of 664 bp in length

* 5328 5427: gap of 100 bp

* 5428 6088: contig of 661 bp in length

* 6089 6188: gap of 100 bp

* 6189 6866: contig of 678 bp in length

* 6867 6966: gap of 100 bp

* 6967 7647: contig of 681 bp in length

* 7648 7747: gap of 100 bp

* 7748 8439: contig of 692 bp in length

* 8440 8539: gap of 100 bp

* 8540 9198: contig of 659 bp in length

* 9199 9298: gap of 100 bp

* 9299 9711: contig of 413 bp in length

* 9712 9811: gap of 100 bp

* 9812 10478: contig of 667 bp in length

* 10479 10578: gap of 100 bp

* 10579 11249: contig of 671 bp in length

* 11250 11349: gap of 100 bp

* 11350 11996: contig of 647 bp in length

* 11997 12096: gap of 100 bp

* 12097 12785: contig of 689 bp in length

* 12786 12885: gap of 100 bp

* 12886 13536: contig of 651 bp in length

* 13537 13636: gap of 100 bp

* 13637 14292: contig of 656 bp in length

* 14293 14392: gap of 100 bp

* 14393 15085: contig of 693 bp in length

* 15086 15185: gap of 100 bp

* 15186 15835: contig of 650 bp in length

* 15836 15935: gap of 100 bp

* 15936 17225: contig of 1290 bp in length

* 17226 17325: gap of 100 bp

* 17326 19181: contig of 1856 bp in length

* 19182 19281: gap of 100 bp

* 19282 21156: contig of 1875 bp in length

* 21157 21256: gap of 100 bp

* 21257 24719: contig of 3463 bp in length

* 24720 24819: gap of 100 bp

* 24820 27090: contig of 2271 bp in length

* 27091 27190: gap of 100 bp

* 27191 30340: contig of 3150 bp in length

* 30341 30440: gap of 100 bp

* 30441 32803: contig of 2363 bp in length

* 32804 32903: gap of 100 bp

* 32904 35837: contig of 2934 bp in length

* 35838 35937: gap of 100 bp

* 35938 39727: contig of 3790 bp in length

* 39728 39827: gap of 100 bp

* 39828 43907: contig of 4080 bp in length

* 43908 44007: gap of 100 bp

* 44008 49223: contig of 5222 bp in length

* 49224 49323: gap of 100 bp

* 49324 52432: contig of 3103 bp in length

* 52433 52532: gap of 100 bp

* 52533 71268: contig of 18736 bp in length

* 71269 71368: gap of 100 bp

* 71369 78519: contig of 7151 bp in length

* 78520 78619: gap of 100 bp

* 78620 87207: contig of 8588 bp in length

* 87208 87307: gap of 100 bp

* 87308 96949: contig of 9642 bp in length

* 96950 97049: gap of 100 bp

* 97050 104314: contig of 7265 bp in length

* 104315 104414: gap of 100 bp

* 104415 113053: contig of 8639 bp in length

* 113054 113153: gap of 100 bp

* 113154 125823: contig of 12670 bp in length

* 125824 125923: gap of 100 bp

* 125924 139598: contig of 13675 bp in length

* 139599 139698: gap of 100 bp

* 139699 153566: contig of 13868 bp in length

* 153567 153666: gap of 100 bp

* 153667 167089: contig of 13423 bp in length

* 167090 167189: gap of 100 bp

* 167190 184611: contig of 17422 bp in length

* 184612 184711: gap of 100 bp

* 184712 205546: contig of 20835 bp in length

* 205547 205646: gap of 100 bp

* 205647 213703: contig of 8057 bp in length.

FEATURES
source

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-157H14"
 /clone_lib="RPCI-23 Female Mouse BAC"

misc_feature
 1..651

misc_feature
 752..1458
 /note="assembly_fragment"
 /note="assembly_fragment"

Query Match 16.0%; Score 33; DB 2; Length 213703;
 Best Local Similarity 50.3%; Pred. No. 16;
 Matches 81; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 4 TCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCGATTAAGGTCCTGTGTCTACTA 63

Db 116461 TCTATCAGGTACTATTTCACATAAATCAACAAACAGACTTTGGTGTGATTACTGTA 116520

QY 64 AGCCAGGTTCTTGTCTTATTATCTTGATTCGTGGCGGTATGTTAAACCCACCTAACCGTT 123

Db 116521 CAGCTGATTGTTTTCACACTACTTGGATGTGAAGGTGTCAATATTGTTGACACAAATATTGTA 116580

QY 124 GTTTGAAGCACTGATTCAGGTATCAAAAAGTGTGT 164
 |||||
 Db 116581 GTTGTGTGATATTTATTTCTAAATTCACAAAAGAGATGT 116621

RESULT 73

AC117425/c

LOCUS

AC117425 143620 bp DNA linear PRI 29-JUN-2002
 Homo sapiens 3 BAC RP11-29718 (Roswell Park Cancer Institute Human
 BAC Library) complete sequence.

AC117425

ACCESSION

AC117425.3

VERSION

HTG.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC117425 143620 bp DNA linear PRI 29-JUN-2002
 Homo sapiens 3 BAC RP11-29718 (Roswell Park Cancer Institute Human
 BAC Library) complete sequence.
 AC117425
 HTG.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 143620)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Albrooks,S.L., Amarantunge,H.C., Are,J.R., Aveye,M., Banks,T.,
 Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Briefe,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Cahay,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.K., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Donahwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gaisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Joudah,S.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,Z., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
 Louisleg,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,B., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
 Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
 Nickerson,E., Nwokenkwo,S., Ogwu,M., Okwuonu,G., Orsanyue,N.,
 Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
 Pickens,R., Primus,E., Pui,L.L., Quiles,M., Ren,Y., Rives,M.,
 Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
 Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
 Sonaite,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
 Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
 Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
 Vera,V., Villalob,B., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
 Warren,R., Washington,C., Watlington,S., Williams,G.,
 Williamson,A., Wleciyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
 Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
 Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 143620)
 Worley,K.C.
 Direct Submission
 Submitted (10-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 143620)
 Worley,K.C.
 Direct Submission
 Submitted (31-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 143620)

Worley,K.C.
 Direct Submission

Submitted (28-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 143620)

Worley,K.C.
 Direct Submission

Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Jun 28, 2002 this sequence version replaced gi:21240517.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
 standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.anno>

ation.html.

FEATURES

source

1. .143620
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-29718"

repeat_region

repeat_region

STS

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

JOURNAL

COMMENT

Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20069615.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GTNB
Center clone name: CH230-242N20
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 117259 bases at least Q40
Consensus quality: 120122 bases at least Q30
Consensus quality: 122747 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1769: contig of 1769 bp in length
1770 1869: gap of unknown length
1870 2963: contig of 1094 bp in length
2964 3063: gap of unknown length
3064 4456: contig of 1393 bp in length
4457 4556: gap of unknown length
4557 5992: contig of 1436 bp in length
5993 6092: gap of unknown length
6093 7489: contig of 1397 bp in length
7490 7589: gap of unknown length
7590 9173: contig of 1584 bp in length
9174 9273: gap of unknown length
9274 11148: contig of 1875 bp in length
11149 13177: contig of 2669 bp in length
13178 14017: gap of unknown length
14018 16203: contig of 2186 bp in length
16204 16303: gap of unknown length
16304 17917: contig of 1614 bp in length
17918 18017: gap of unknown length
18018 19705: contig of 1688 bp in length
19706 22857: contig of 3052 bp in length
22858 22957: gap of unknown length
22959 25118: contig of 2161 bp in length
25119 27161: contig of 1943 bp in length
27162 27261: gap of unknown length
27262 29591: contig of 2330 bp in length
29592 30909: contig of 1218 bp in length
30910 31009: gap of unknown length
31010 33313: contig of 2304 bp in length
33314 33414: gap of unknown length
33415 36514: contig of 3001 bp in length
36515 41632: contig of 5118 bp in length
41633 45404: contig of 3672 bp in length
45405 45504: gap of unknown length
45505 50390: contig of 4886 bp in length
50391 50490: gap of unknown length
50491 54730: contig of 4240 bp in length

Db 50980 ATGTACAGTAACTTACTGAACTAGCTTGATGCTGCTAGTGTCTTAAGTCTGT 51039
QY 191 CTTTCCA 197
Db 51040 CTTTACA 51046

RESULT 75
AC117109
LOCUS
DEFINITION Rattus norvegicus clone CH230-242N20, *** SEQUENCING IN PROGRESS
***, 34 unordered pieces.
AC117109
VERSION AC117109.2 GI:21745874
KEYWORDS HTG, HTGS, PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 142597)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbaria, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Fallis, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Homi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, J., Johnson, R., Jollivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oraguve, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 142597)
Worley, K.C.
Direct Submission
TITLE
Submitted (08-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 142597)
Worley, K.C.
Direct Submission
REFERENCE
AUTHORS
TITLE

```

* 54731 54830: gap of unknown length
* 54831 58409: contig of 3579 bp in length
* 58410 58509: gap of unknown length
* 58510 63417: contig of 4908 bp in length
* 63418 63517: gap of unknown length
* 63518 68861: contig of 5344 bp in length
* 68862 68961: gap of unknown length
* 68962 74430: contig of 5469 bp in length
* 74431 74530: gap of unknown length
* 74531 81012: contig of 6482 bp in length
* 81013 81112: gap of unknown length
* 81113 88161: contig of 7049 bp in length
* 88162 88261: gap of unknown length
* 88262 94871: contig of 6610 bp in length
* 94872 94972: gap of unknown length
* 94973 103462: contig of 8491 bp in length
* 103463 103562: gap of unknown length
* 103563 111682: contig of 8120 bp in length
* 111683 111782: gap of unknown length
* 111783 119258: contig of 7476 bp in length
* 119259 119358: gap of unknown length
* 119359 129288: contig of 9930 bp in length
* 129289 129388: gap of unknown length
* 129389 142597: contig of 13209 bp in length.

```

FEATURES

```

Location/Qualifiers
1..142597
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-242N20"

```

```

BASE COUNT 40398 a 27668 c 27321 g 40112 t 7098 others
ORIGIN

```

```

Query Match 15.8%; Score 32.6; DB 2; Length 142597;
Best Local Similarity 52.6%; Pred. NO. 21;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 7 AGCTGGTACCATACCTGCATATGCTCAGAACACAGTTAAAGTCTCTGTCTACTAAGC 66
Db 120419 AGCTCTCTCCCTCCTCCTCGATTTCGAAATTCACACTACGAGCATGTGAGACCCCTGC 120478
QY 67 CAGGTTCTTGTCTCTATTATCTTGTGCTTGTGCTTGTAAACCCACCTAACCGTTGTT 126
Db 120479 CTTCTCTCTTTTGAATTATGTTGATGAATTGACAGAAGTGGAGGCCCATGCTACAGATCT 120538
QY 127 TGAAGACACTGATT 141
Db 120539 TAAATCCTTGGATT 120553

```

```

Search completed: February 15, 2003, 22:44:04
Job time : 1527.5 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 20:20:11 ; Search time 66 Seconds
(without alignments)
264.857 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvtkgscpi1.....cpgikkcccgscgmactvfpq 57

Scoring table: BLOSUM62

Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09833799/runat_12022003_120043_23688/app_query.fasta_1.199
-DB=Issued_Patents.NA -QPM=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=100 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799 @CGN 1.1 61 @runat_12022003_120043_23688 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NSG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WASN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	571	1	US-08-322-742-14
2	324	99.1	177	1	US-08-379-437-3
3	323	98.8	177	1	US-08-379-437-5
4	323	98.8	177	1	US-08-379-437-7
5	322	98.5	321	1	US-08-322-742-11
6	143	43.7	134	2	US-07-963-538B-5
7	139	42.5	180	3	US-08-483-503A-2
8	105	32.1	292	4	US-09-404-879A-377
9	98	30.0	724	4	US-09-724-864-23
10	94.5	28.9	437	4	US-09-383-586-29
11	92.5	28.3	6314	1	US-08-211-430-1
12	85	26.0	753	4	US-09-152-060-34

13	85	26.0	783	4	US-09-152-060-50	Sequence 50, Appl
14	81.5	24.9	1042	2	US-08-761-248B-5	Sequence 5, Appl
15	81.5	24.9	1071	2	US-08-761-248B-1	Sequence 1, Appl
16	79.5	24.3	1366	2	US-08-761-248B-3	Sequence 3, Appl
17	77.5	23.7	4157	2	US-08-162-146-2	Sequence 2, Appl
18	77.5	23.7	4157	4	US-09-314-127-2	Sequence 2, Appl
19	75	22.9	80	1	US-08-304-051-7	Sequence 7, Appl
20	75	22.9	80	1	US-08-304-051-11	Sequence 11, Appl
21	75	22.9	80	5	PCT-US95-11445-7	Sequence 7, Appl
22	75	22.9	80	5	PCT-US95-11445-11	Sequence 11, Appl
23	69	21.1	403	4	US-09-404-879A-198	Sequence 198, App
24	69	21.1	2065	3	US-08-335-865J-8	Sequence 8, Appl
25	68	20.8	1419	4	US-03-434-288-2	Sequence 2, Appl
26	66.5	20.3	11444	2	US-08-222-617A-26	Sequence 26, Appl
27	66.5	20.3	12364	2	US-08-222-617A-1	Sequence 1, Appl
28	65	19.9	68	2	US-07-963-538B-34	Sequence 34, Appl
29	64.5	19.7	1364	1	US-08-306-691B-50	Sequence 50, Appl
30	64.5	19.7	1364	5	PCT-US93-06251-65	Sequence 65, Appl
31	64.5	19.7	1590	4	US-08-961-527-184	Sequence 184, App
32	64.5	19.7	2625	2	US-08-357-533A-1	Sequence 1, Appl
33	64.5	19.7	2625	2	US-08-459-009-1	Sequence 1, Appl
34	64.5	19.7	2625	3	US-08-459-951-1	Sequence 1, Appl
35	64	19.6	80161	3	US-09-036-987A-1	Sequence 1, Appl
36	64	19.6	80161	4	US-03-370-700-1	Sequence 1, Appl
37	63.5	19.4	4170	2	US-08-619-198-4	Sequence 4, Appl
38	63.5	19.4	5192	2	US-08-619-198-8	Sequence 8, Appl
39	63.5	19.4	5198	2	US-08-619-198-1	Sequence 1, Appl
40	63	19.3	8931	3	US-09-028-934-28	Sequence 28, Appl
41	62.5	19.1	3855	4	US-08-974-549A-4	Sequence 4, Appl
42	62.5	19.1	4015	3	US-08-851-843A-224	Sequence 224, App
43	62.5	19.1	4015	4	US-08-974-549A-1	Sequence 1, Appl
44	62.5	19.1	4015	4	US-08-854-050-224	Sequence 224, App
45	62.5	19.1	4015	4	US-09-430-323-224	Sequence 224, App
46	62.5	19.1	4015	4	US-09-572-423B-3	Sequence 3, Appl
47	62.5	19.1	4015	4	US-09-128-354-1	Sequence 1, Appl
48	62.5	19.1	4015	4	US-09-675-321-1	Sequence 1, Appl
49	62.5	19.1	4015	4	US-09-052-919-1	Sequence 1, Appl
50	62.5	19.1	4037	4	US-08-974-549A-343	Sequence 343, App
51	62.5	19.1	4335	4	US-08-974-549A-6	Sequence 6, Appl
52	62.5	19.1	6677	4	US-08-939-366-27	Sequence 27, Appl
53	62.5	19.1	6677	4	US-08-467-997-6	Sequence 6, Appl
54	61.5	18.8	708	2	US-08-400-159-3	Sequence 3, Appl
55	61.5	18.8	708	3	US-08-611-729A-3	Sequence 3, Appl
56	61.5	18.8	1665	2	US-08-886-152-2	Sequence 2, Appl
57	61.5	18.8	1665	4	US-09-196-222-2	Sequence 2, Appl
58	61.5	18.8	2892	1	US-08-264-534-5	Sequence 5, Appl
59	61.5	18.8	2892	1	US-08-083-590A-1	Sequence 1, Appl
60	61.5	18.8	2892	2	US-08-465-500-5	Sequence 5, Appl
61	61.5	18.8	2892	2	US-08-346-126-5	Sequence 5, Appl
62	61.5	18.8	2892	2	US-08-346-128-5	Sequence 5, Appl
63	61.5	18.8	2892	3	US-08-532-384-1	Sequence 1, Appl
64	61.5	18.8	2892	3	US-08-893-828-5	Sequence 5, Appl
65	61	18.7	631	2	US-08-747-887-4	Sequence 4, Appl
66	61	18.7	631	2	US-08-728-323A-4	Sequence 4, Appl
67	61	18.7	653	4	US-08-998-416-149	Sequence 149, App
68	61	18.7	1055	4	US-09-056-556-191	Sequence 191, App
69	61	18.7	1055	4	US-09-072-596-186	Sequence 186, App
70	60	18.3	518	4	US-09-615-192A-172	Sequence 172, App
71	60	18.3	850	4	US-09-449-285A-8	Sequence 8, Appl
72	60	18.3	1089	4	US-08-858-207A-161	Sequence 161, App
73	60	18.3	2783	4	US-09-282-147-37	Sequence 37, Appl
74	60	18.3	3222	4	US-09-543-084A-29	Sequence 29, Appl
75	60	18.3	4050	4	US-09-543-084A-26	Sequence 26, Appl
76	60	18.3	4093	4	US-09-543-084A-28	Sequence 28, Appl
77	60	18.3	8802	3	US-08-896-449A-1	Sequence 1, Appl
78	60	18.3	8802	3	US-09-132-652-1	Sequence 1, Appl
79	60	18.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl
80	60	18.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl
81	59.5	18.2	703	4	US-09-452-817-5	Sequence 5, Appl
82	59.5	18.2	717	4	US-08-998-416-849	Sequence 849, App
83	59.5	18.2	780	4	US-09-452-817-7	Sequence 7, Appl
84	59	18.0	67	2	US-07-963-538B-33	Sequence 33, Appl
85	59	18.0	610	4	US-09-072-596-293	Sequence 293, App

Sequence 2, Appli
Sequence 36, Appli
Sequence 2, Appli
Sequence 122, App
Sequence 15, Appli
Sequence 15, Appli
Sequence 347, App
Sequence 4, Appli
Sequence 175, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli

86 59 18.0 2912 5 PCT-US91-01726-2
87 59 18.0 6719 4 US-09-740-235-36
88 59 18.0 8021 4 US-09-740-235-2
89 58.5 17.9 428 1 US-08-470-179-122
90 58.5 17.9 440 4 US-08-821-827C-15
91 58.5 17.9 440 4 US-09-290-202B-15
92 58.5 17.9 1740 4 US-09-643-597-347
93 58.5 17.9 4159 4 US-09-061-709-4
94 58.5 17.9 4181 4 US-09-643-597-175
95 58.5 17.9 7001 1 US-08-258-261B-1
96 58.5 17.9 7001 1 US-08-456-837-1
97 58.5 17.9 7001 1 US-08-457-342-1
98 58.5 17.9 7001 1 US-08-457-646A-1
99 58.5 17.9 7001 1 US-08-458-076A-1
100 58.5 17.9 7001 1 US-08-457-335A-1

ALIGNMENTS

RESULT 1

US-08-322-742-14
; Sequence 14, Application US/08322742
; Patent No. 568641
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/322,742
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: September 1, 1992
; APPLICATION NUMBER: 07/844,296
; FILING DATE: February 28, 1992
; APPLICATION NUMBER: 07/552,216
; FILING DATE: February 28, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/048003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 14:
; . SEQUENCE CHARACTERISTICS:
; LENGTH: 571
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

US-08-322-742-14
Alignment Scores:
Pred. No.: 2.1e-33 Length: 571
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 1

09-833799-13B (1-57) x US-08-322-742-14 (1-571)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 199 GGCAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCCATATCTTG 258
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 259 ATCCGGTGGCCCATGTTGAATCTCTCTTAACCGCTGCTTGAAGATACTGACTGCCACGA 318
Qy 41 IleLysLysCysGlyGlySerCysGlyMetAlaCysPheValProGln 57
Db 319 ATCAAGAAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCCAG 369

RESULT 2

US-08-379-437-3
; Sequence 3, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
; APPLICANT: ISHIMA, YOSHIKI
; APPLICANT: OKAWA, NORIYUKI
; APPLICANT: YOSHIDA, MASAYA
; APPLICANT: AMAGAYA, SAKAE
; APPLICANT: KAJI, AKIRA
; TITLE OF INVENTION: NOVEL ELAFIN DERIVATIVE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,437
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP93/01133
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 234085/1992
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..171
; US-08-379-437-3

Alignment Scores:
Pred. No.: 1.05e-33 Length: 177
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0


```
Query Match: 99.08% Indels: 0
DB: 1 Gaps: 0
09-833799-13B (1-57) x US-08-379-437-3 (1-177)
Qy 1 AlaGlnGluProValIysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
Db 1 GCACAGGAACACAGTTAAAGGTCCGGTGTGACCAACCGGGCTCTTCCCGATTATCCTG 60
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuIysAspThrAspCysProGly 40
Db 61 ATCCGCTCGCTTTCCTGAACCCCGAACCGTTGTCTGAAGACACTGACTGCCCGGGT 120
Qy 41 IleIysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAATGCTGCAAGGTCTTGGCGTATGCGCATGCTTCGTTCCGCAG 171
RESULT 3
US-08-379-437-5
; Sequence 5, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
; APPLICANT: ISHIMA, YOSHIAKI
; APPLICANT: OKAWA, NORIYUKI
; APPLICANT: YOSHIDA, MASAYA
; APPLICANT: AMAGAYA, SAKAE
; APPLICANT: KAJI, AKIRA
; TITLE OF INVENTION: NOVEL ELAFIN DERIVATIVE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,437
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP93/01133
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 234085/1992
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..171
US-08-379-437-5
Alignment Scores: 1.41e-33 Length: 177
Pred. No.: 1.41e-33 Length: 177

Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
09-833799-13B (1-57) x US-08-379-437-5 (1-177)
Qy 1 AlaGlnGluProValIysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
Db 1 GCACAGGAACACAGTTAAAGGTCCGGTGTGACCAACCGGGCTCTTCCCGATTATCCTG 60
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuIysAspThrAspCysProGly 40
Db 61 ATCCGCTCGCTTTCCTGAACCCCGAACCGTTGTCTGAAGACACTGACTGCCCGGGT 120
Qy 41 IleIysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAATGCTGCAAGGTCTTGGCGTATGCGCATGCTTCGTTCCGCAG 171
RESULT 4
US-08-379-437-7
; Sequence 7, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
; APPLICANT: ISHIMA, YOSHIAKI
; APPLICANT: OKAWA, NORIYUKI
; APPLICANT: YOSHIDA, MASAYA
; APPLICANT: AMAGAYA, SAKAE
; APPLICANT: KAJI, AKIRA
; TITLE OF INVENTION: NOVEL ELAFIN DERIVATIVE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,437
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP93/01133
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 234085/1992
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..171
US-08-379-437-7
```

Alignment Scores:

Pred. No.: 1.41e-33 Length: 177
 Score: 323.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.25% Mismatches: 0
 Query Match: 98.78% Indels: 0
 DB: 1 Gaps: 0

09-833799-13B (1-57) x US-08-379-437-7 (1-177)

QY 1 AlaGlnGluProValIysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
 DB 1 GCACAGGAACCCAGTAAAGGTCGGGTGCGACCAACCGGGCTCTTGCCCGATTATCCTG 60
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIysAspThrAspCysProGly 40
 DB 61 ATCCGCTGGCTATTCTGAACCCGCGAACCGTCTCTGAAAGACACTGACTGCCCGGT 120
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 DB 121 ATCAAAAAATGCTGCGAAGTTCCTTGGCGGTATGCGATGCTTGGTTCGCGCAG 171

RESULT 5

US-08-322-742-11

; Sequence 11, Application US/08322742

; Patent No. 5688641

; GENERAL INFORMATION:

; APPLICANT: Sager, Ruth

; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM PS/2 Model 502 or 55SX

; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/322,742

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/938,823

; FILING DATE: September 1, 1992

; APPLICATION NUMBER: 07/844,296

; FILING DATE: February 28, 1992

; APPLICATION NUMBER: 07/552,216

; FILING DATE: February 28, 1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fraser, Janis K.

; REGISTRATION NUMBER: 34,819

; REFERENCE/DOCKET NUMBER: 00530/048003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: (617) 542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 321

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; US-08-322-742-11

Alignment Scores:

Pred. No.: 4.23e-33 Length: 321
 Score: 322.00 Matches: 56
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.47% Indels: 0
 DB: 1 Gaps: 0

09-833799-13B (1-57) x US-08-322-742-11 (1-321)

QY 1 AlaGlnGluProValIysGlyProValSerThrIysProGlySerCysProIleIleLeu 20
 DB 144 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCCTGCTCTGCCCAATTATCTTG 203
 QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuIysAspThrAspCysProGly 40
 DB 204 ATCCGTCGCGCATGTGAATCTCTTAACCGTGTCTTGAAGATACACTGCGCCAGGA 263
 QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 DB 264 ATCAAGAAGTGTGTAAGGCTCTTGGGGATGGCTGTTCGTTCCCTCCC 311

RESULT 6

US-07-963-538B-5

; Sequence 5, Application US/07963538B

; Patent No. 5851983

; GENERAL INFORMATION:

; APPLICANT: SUGIYAMA, TAKASHI

; APPLICANT: KAMIMURA, TAKASHI

; APPLICANT: MASUDA, KENICHI

; APPLICANT: OKADA, MASAHIRO

; APPLICANT: OHTSUKA, EIKO

; APPLICANT: IMAIZUMI, ATSUSHI

; APPLICANT: WATANABE, KUNIHITO

; APPLICANT: SUGA, TETSUYA

; APPLICANT: MATSUMOTO, YOHICHI

; APPLICANT: TAKEUCHI, AKIKO

; TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND

; TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: COOLEY GODWARD LLP

; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR

; CITY: PALO ALTO

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-2155

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/963,538B

; FILING DATE: 20-OCT-1992

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/843,359

; FILING DATE: 25-FEB-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/408,483

; FILING DATE: 22-AUG-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 4-212399

; FILING DATE: 17-JUL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 4-212398

; FILING DATE: 17-JUL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 3-355553

; FILING DATE: 24-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 62-330219

; FILING DATE: 28-DEC-1987

; ATTORNEY/AGENT INFORMATION:

; NAME: NEELEY PH.D., RICHARD L.

Db 61 TTGAAGTGTTCATGGGCATGTGTGGAAATCCTCGTTCCCT 105

RESULT 9

US-09-724-864-23
 ; Sequence 23, Application US/09724864
 ; Patent No. 6380362
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Murison, James G.
 ; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
 ; by the polynucleotides and methods for their use.
 ; FILE REFERENCE: 11000.105001
 ; CURRENT APPLICATION NUMBER: US/09/724,864
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 23
 ; LENGTH: 724
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-09-724-864-23

Alignment Scores:
 Pred. No.: 0.000767 Length: 724
 Score: 98.00 Matches: 25
 Percent Similarity: 43.24% Conservative: 7
 Best Local Similarity: 33.78% Mismatches: 18
 Query Match: 29.97% Indels: 24
 DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-724-864-23 (1-724)

QY 3 GluProValLysGlyProValSerThrLysPro-----13
 Db 313 AAACACCGGGAGGTCAAGTCTCCAGAACCCCGGCTGTGACAGGAGGCTTAGGT 372
 QY 14 -----GlySerCysProIlelle-----LeuIleArgCysala 24
 Db 373 GTCCGAGAAAGCAGGCGCCTGCCCCAGCGTGGACATACCCAGCTCGGCTCTGT--- 429
 QY 25 MetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysCys 44
 Db 430 -----GAGGACCAAGTGTGAGTGGACAGCCAGTGTTCGGCAACATGAATGC 477
 QY 45 CysGluGlySerCysGly---MetAlaCysPheValProGln 57
 Db 478 TGCCGCAATGGATGTGGGAAGATGGCTGCACCAACCCAAA 519

RESULT 10

US-09-383-586-29
 ; Sequence 29, Application US/09383586
 ; Patent No. 6242419
 ; GENERAL INFORMATION:
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Abernethy, Nevin
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Kumble, Anand
 ; APPLICANT: Murison, Greg
 ; TITLE OF INVENTION: Compounds isolated from stromal cells
 ; FILE REFERENCE: 11000.1037c1
 ; CURRENT APPLICATION NUMBER: US/09/383,586
 ; CURRENT FILING DATE: 1999-08-26
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 29
 ; LENGTH: 437
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-09-383-586-29

Alignment Scores:

Pred. No.: 0.00109 Length: 437
 Score: 94.50 Matches: 19
 Percent Similarity: 41.51% Conservative: 3
 Best Local Similarity: 35.85% Mismatches: 14
 Query Match: 28.90% Indels: 17
 DB: 4 Gaps: 2

09-833799-13B (1-57) x US-09-383-586-29 (1-437)

QY 12 LysProGlySerCysProIlelleLeuIleArgCysAlaMetLeuAsnProPro-----29
 Db 80 AAACCTGGAGCTGTGCC-----AAGCCTTCACCCAGAA 112
 QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
 Db 113 AGTGTGGGAATTTGTTGATCAATGCTCAGGAGATGGATCCTGCCCTGGCAACATGAAG 172
 QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 Db 173 TGCTGTAGCAATAGCTGTGTCATGCTCTGCAAAACTCCT 211

RESULT 11

US-08-211-430-1
 ; Sequence 1, Application US/08211430
 ; Patent No. 5763166
 ; GENERAL INFORMATION:
 ; APPLICANT:

;; TITLE OF INVENTION: NUCLEIC SEQUENCE OF THE GENE ASSOCIATED WITH
 ;; TITLE OF INVENTION: X CHROMOSOME LINKED KALLMANN SYNDROME, CORRESPONDING
 ;; TITLE OF INVENTION: PEPTIDE SEQUENCES, DIAGNOSTIC APPLICATIONS.
 ;; NUMBER OF SEQUENCES: 32
 ;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
 ;; COMPUTER: IBM PC compatible
 ;; OPERATING SYSTEM: PC-DOS/MS-DOS
 ;; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ;; CURRENT APPLICATION DATA:
 ;; APPLICATION NUMBER: US/08/211,430
 ;; INFORMATION FOR SEQ ID NO: 1:
 ;; SEQUENCE CHARACTERISTICS:
 ;; LENGTH: 6314 base pairs
 ;; TYPE: nucleic acid
 ;; STRANDEDNESS: double
 ;; TOPOLOGY: linear
 ;; MOLECULE TYPE: cDNA
 ;; HYPOTHETICAL: NO
 ;; ANTI-SENSE: NO
 ;; ORIGINAL SOURCE:
 ;; ORGANISM: Homo sapiens
 ;; CELL LINE: foetal brain cells
 US-08-211-430-1

Alignment Scores:

Pred. No.: 0.0726 Length: 6314
 Score: 92.50 Matches: 18
 Percent Similarity: 40.38% Conservative: 3
 Best Local Similarity: 34.62% Mismatches: 20
 Query Match: 28.29% Indels: 11
 DB: 1 Gaps: 2

09-833799-13B (1-57) x US-08-211-430-1 (1-6314)

QY 12 LysProGlySerCysProIlelle-----IleLeuIleArgCysAlaMet 25
 Db 538 AACGAGGGGAGCTGTCCGGCTCCTGAGAAAGCAGTGGATTTTGGCGCCGCTGTGT--- 594
 QY 26 LeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysCysCys 45
 Db 595 -----GAAAGCTGCCAAGTTGACATGCTCTGGGGTGAAGAAATGTTGT 642
 QY 46 GluGlySerCysGlyMetAlaCysPheValProGln 57


```
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,348
; FILING DATE: 07-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Turley, Charles P
; REGISTRATION NUMBER: 35,723
; REFERENCE/DOCKET NUMBER: 34012.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)9513310
; TELEFAX: (713)9513314
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1071 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-761-248B-5
;
; Alignment Scores:
; Pred. No.: 0.165 Length: 1042
; Score: 81.50 Matches: 14
; Percent Similarity: 43.48% Conservative: 6
; Best Local Similarity: 30.43% Mismatches: 15
; Query Match: 24.92% Indels: 11
; DB: 2 Gaps: 1
;
; 09-833799-13B (1-57) x US-08-761-248B-5 (1-1042)
;
; QY 8 ProValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn 27
;
; Db 235 CCACGACGCTACCCCGGGTGCCTGT----- 261
;
; QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
;
; Db 262 CAGGCCACACGCTGCAGTCTGACTCTGAGTGCACACGACGACGCTGTTTACAAC 321
;
; QY 48 SerCysGlyMetAlaCys 53
;
; Db 322 GGCTGTGCCTATGCCTGC 339
;
; RESULT 15
; US-08-761-248B-1
; Sequence 1, Application US/08761248B
; Patent No. 5958735
; GENERAL INFORMATION:
; APPLICANT: ROWLEY,, DAVID R.
; TITLE OF INVENTION: UROGENITAL SINUS DERIVED GROWTH
; FACTOR NUCLEOTIDE AND AMINO ACID SEQUENCES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jenkins & Gilchrist
; STREET: 1100 Louisiana, Suite 1800
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,248B
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,348
; FILING DATE: 07-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Turley, Charles P
; REGISTRATION NUMBER: 35,723
; REFERENCE/DOCKET NUMBER: 34012.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)9513310
; TELEFAX: (713)9513314
;
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,348
; FILING DATE: 07-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Turley, Charles P
; REGISTRATION NUMBER: 35,723
; REFERENCE/DOCKET NUMBER: 34012.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)9513310
; TELEFAX: (713)9513314
;
; NAME: Turley, Charles P
; REGISTRATION NUMBER: 35,723
; REFERENCE/DOCKET NUMBER: 34012.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)9513310
; TELEFAX: (713)9513314
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1071 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-761-248B-1
;
; Alignment Scores:
; Pred. No.: 0.172 Length: 1071
; Score: 81.50 Matches: 14
; Percent Similarity: 43.48% Conservative: 6
; Best Local Similarity: 30.43% Mismatches: 15
; Query Match: 24.92% Indels: 11
; DB: 2 Gaps: 1
;
; 09-833799-13B (1-57) x US-08-761-248B-1 (1-1071)
;
; QY 8 ProValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn 27
;
; Db 235 CCACGACGCTACCCCGGGTGCCTGT----- 261
;
; QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
;
; Db 262 CAGGCCACACGCTGCAGTCTGACTCTGAGTGCACACGACGACGCTGTTTACAAC 321
;
; QY 48 SerCysGlyMetAlaCys 53
;
; Db 322 GGCTGTGCCTATGCCTGC 339
;
; RESULT 16
; US-08-761-248B-3
; Sequence 3, Application US/08761248B
; Patent No. 5958735
; GENERAL INFORMATION:
; APPLICANT: ROWLEY,, DAVID R.
; TITLE OF INVENTION: UROGENITAL SINUS DERIVED GROWTH
; FACTOR NUCLEOTIDE AND AMINO ACID SEQUENCES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jenkins & Gilchrist
; STREET: 1100 Louisiana, Suite 1800
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,248B
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,348
; FILING DATE: 07-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Turley, Charles P
; REGISTRATION NUMBER: 35,723
; REFERENCE/DOCKET NUMBER: 34012.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)9513310
; TELEFAX: (713)9513314
```

```

;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1366 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-761-248B-3

Alignment Scores:
Pred. No.: 0.431 Length: 1366
Score: 79.50 Matches: 14
Percent Similarity: 43.48% Conservative: 6
Best Local Similarity: 30.43% Mismatches: 15
Query Match: 24.31% Indels: 11
DB: 2 Gaps: 1

09-833799-13B (1-57) x US-08-761-248B-3 (1-1366)
QY 8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn 27
Db 362 CCGGGAGCGTCCCGCGCGCGCTGCCAGCGCGG----- 397
QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysCysCysGluGly 47
Db 398 -----CGCTGTCAAGCGGACTCGAGTGCCTCGCGCGGCGCGCGCTGCTGTACAAC 448
QY 48 SerCysGlyMetAlaCys 53
Db 449 GGATGCGCTACGCGCTGC 466

RESULT 17
US-08-162-146-2
; Sequence 2, Application US/08162146
; Patent No. 5965788
; GENERAL INFORMATION:
; APPLICANT: HOUEBINE, Louis-Marie
; APPLICANT: DEVINOT, Eve
; APPLICANT: THEPOT, Dominique
; TITLE OF INVENTION: Production of a Protein of Interest in
; TITLE OF INVENTION: the Milk of a Transgenic Mammalian
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,146
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR92/00533
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91/07179
; FILING DATE: 12-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1868..1949, 2462..2587, 2888..3046, 3416
; LOCATION: ..3429)
US-08-162-146-2

Alignment Scores:
Pred. No.: 3.5 Length: 4157
Score: 77.50 Matches: 20
Percent Similarity: 46.30% Conservative: 5
Best Local Similarity: 37.04% Mismatches: 22
Query Match: 23.70% Indels: 7
DB: 2 Gaps: 3

09-833799-13B (1-57) x US-08-162-146-2 (1-4157)
QY 8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeu 26
Db 2884 CAGTCCCTACCCCAAGGCTGGCGCTGCC-----TGGGTGCAGCGCCCAATGCTG 2937
QY 27 Asn-----ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLys 42
Db 2938 TCCCGATTGTGTGAGGAGCTGAGCGACTGTGCCAACGACATCGAGTCGAGGGCGCAAG 2997
QY 43 LysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 2998 AAGTGTCTTTTCAGCGCGCTGCCCATCGCTATCTGGAACCC 3039

RESULT 18
US-09-314-127-2
; Sequence 2, Application US/09314127
; Patent No. 6268545
; GENERAL INFORMATION:
; APPLICANT: HOUEBINE, Louis-Marie
; APPLICANT: DEVINOT, Eve
; APPLICANT: THEPOT, Dominique
; TITLE OF INVENTION: Production of a Protein of Interest in
; TITLE OF INVENTION: the Milk of a Transgenic Mammalian
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,127
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/162,146
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91/07179
; FILING DATE: 12-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(1868..1949, 2462..2587, 2888..3046, 3416
; LOCATION: ..3429)
US-09-314-127-2

Alignment Scores:
Pred. No.: 3.5 Length: 4157
Score: 77.50 Matches: 20
Percent Similarity: 46.30% Conservative: 5
Best Local Similarity: 37.04% Mismatches: 22
Query Match: 23.70% Indels: 7
DB: 4 Gaps: 3

09-833799-13B (1-57) x US-09-314-127-2 (1-4157)

QY 8 ProValSerThr---LysProGlySerCysProIleLeuLeuArgCysAlaMetLeu 26
Db 2884 CCAATCCCTACCCCAAGGCTGCCGCTGCCCC-----TGGGTGCAGCGGCCCAATGCTG 2937
QY 27 Asn-----ProProAnArgCysLeuLysAspThrAspCysProGlyIleLys 42
Db 2938 TCCAGTGTGTGAGGAGCTGAGCGACTGTGCCAAGCATCATGAGTGCAGGGCGCACAG 2997
QY 43 LysCysCysGluCysGlyMetAlaCysPheValPro 56
Db 2998 AAGTGTCTTACGCGCTGCGCCATCGCTATCTGGAACCC 3039

RESULT 19

US-08-304-051-7
; Sequence 7, Application US/08304051

; Patent No. 5633227
; GENERAL INFORMATION:
; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.
; TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
; INHIBITOR OF TRYPTASE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Miles Inc.
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word Perfect 3.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,051
; FILING DATE: 12-Sept-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA: No. 5633227e
; ATTORNEY/AGENT INFORMATION:
; NAME: William F. Gray
; REGISTRATION NUMBER: 31018
; REFERENCE/DOCKET NUMBER: MWH 322

TELEPHONE: (203) 937-2712
TELEFAX: (203) 937-2795
TELEX: 221949 MILES UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 bases
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: - Other nucleic acid: synthetic
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FEATURE:
; NAME/KEY: Primer for Polymerase Chain Reaction used to make mutin
; NAME/KEY: of human SLPI.
US-08-304-051-7

Alignment Scores:
Pred. No.: 0.0354 Length: 80
Score: 75.00 Matches: 12
Percent Similarity: 73.68% Conservative: 2
Best Local Similarity: 63.16% Mismatches: 5
Query Match: 22.94% Indels: 0
DB: 1 Gaps: 0

09-833799-13B (1-57) x US-08-304-051-7 (1-80)

QY 12 LysProGlySerCysProIleLeuLeuArgCysAlaMetLeuAsnProAsn 30
Db 21 AAGCTGGGAAGTGGCCAGTACTTATGCGCAATGATGATGCTTAACCCCCCAAT 77

RESULT 20

US-08-304-051-11

; Sequence 11, Application US/08304051

; Patent No. 5633227

; GENERAL INFORMATION:

; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.
; TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
; INHIBITOR OF TRYPTASE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Miles Inc.
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word Perfect 3.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,051
; FILING DATE: 12-Sept-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA: No. 5633227e
; ATTORNEY/AGENT INFORMATION:
; NAME: William F. Gray
; REGISTRATION NUMBER: 31018
; REFERENCE/DOCKET NUMBER: MWH 322

TELEPHONE: (203) 937-2712
TELEFAX: (203) 937-2795
TELEX: 221949 MILES UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; MOLECULE TYPE: - Other nucleic acid: synthetic
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FEATURE:
; NAME/KEY: Primer for Polymerase Chain Reaction used to make mutin
; NAME/KEY: of human SLPI.
US-08-304-051-11

Alignment Scores:

Pred. No.: 0.0354 Length: 80
Score: 75.00 Matches: 12
Percent Similarity: 73.68% Conservative: 2
Best Local Similarity: 63.16% Mismatches: 5
Query Match: 22.94% Indels: 0
DB: 1 Gaps: 0

09-833799-13B (1-57) x US-08-304-051-11 (1-80)

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsn 30
Db 21 AAGCTGGGAAGTGGCCAGTGACTTATGGCCAATGTAAGATGCTTAACCCCAAT 77

RESULT 21

PCT-US95-11445-7

; Sequence 7, Application PC/TUS9511445

; GENERAL INFORMATION:

; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria,

; APPLICANT: Katherine A.

; TITLE OF INVENTION: Secretory Leukocyte Protease

; TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bayer Corporation

; STREET: 400 Morgan Lane

; CITY: West Haven

; STATE: Connecticut

; COUNTRY: USA

; ZIP: 06516

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: System 7.1

; SOFTWARE: Word Perfect 3.0a

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/11445

; FILING DATE: 11 September 1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/304,051

; FILING DATE: 12 SEPTEMBER 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: William F. Gray

; REGISTRATION NUMBER: 31018

; REFERENCE/DOCKET NUMBER: MWH 322P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (203) 937-2712

; TELEFAX: (203) 937-5492

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 80 bases

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE:

; MOLECULE TYPE: - Other nucleic Acid: synthetic

; HYPOTHETICAL: No

; ANTI-SENSE: No

; FEATURE:

; NAME/KEY: Primer for Polymerase Chain Reaction used to

; NAME/KEY: make mutein of human SLPI.

PCT-US95-11445-7

Alignment Scores:

Pred. No.: 0.0354 Length: 80
Score: 75.00 Matches: 12
Percent Similarity: 73.68% Conservative: 2
Best Local Similarity: 63.16% Mismatches: 5
Query Match: 22.94% Indels: 0
DB: 5 Gaps: 0

09-833799-13B (1-57) x PCT-US95-11445-7 (1-80)

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsn 30
Db 21 AAGCTGGGAAGTGGCCAGTGACTTATGGCCAATGTAAGATGCTTAACCCCAAT 77

RESULT 22

PCT-US95-11445-11

; Sequence 11, Application PC/TUS9511445

; GENERAL INFORMATION:

; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria,

; APPLICANT: Katherine A.

; TITLE OF INVENTION: Secretory Leukocyte Protease

; TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bayer Corporation

; STREET: 400 Morgan Lane

; CITY: West Haven

; STATE: Connecticut

; COUNTRY: USA

; ZIP: 06516

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: System 7.1

; SOFTWARE: Word Perfect 3.0a

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/11445

; FILING DATE: 11 September 1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/304,051

; FILING DATE: 12 SEPTEMBER 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: William F. Gray

; REGISTRATION NUMBER: 31018

; REFERENCE/DOCKET NUMBER: MWH 322P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (203) 937-2712

; TELEFAX: (203) 937-5492

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 80 bases

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE:

; MOLECULE TYPE: - Other nucleic Acid: synthetic

; HYPOTHETICAL: No

; ANTI-SENSE: No

; FEATURE:

; NAME/KEY: Primer for Polymerase Chain Reaction used to

; NAME/KEY: make mutein of human SLPI.

PCT-US95-11445-11

Alignment Scores:

Pred. No.: 0.0354 Length: 80
Score: 75.00 Matches: 12
Percent Similarity: 73.68% Conservative: 2
Best Local Similarity: 63.16% Mismatches: 5
Query Match: 22.94% Indels: 0
DB: 5 Gaps: 0

09-833799-13B (1-57) x PCT-US95-11445-11 (1-80)

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsn 30
Db 21 AAGCTGGGAAGTGGCCAGTGACTTATGGCCAATGTAAGATGCTTAACCCCAAT 77

RESULT 23

US-09-404-879A-198/c

; Sequence 198, Application US/09404879A

; Patent No. 6468546

QY 32 -----CysLeuLysAspThrAsp---CysProGlyIleLys----- 42
Db 846 GTCCCGTGGCGCTCGACGACATCGACCTCGCGCGGTGACGGGGGTGGCC 905
QY 43 -----LysCysCysGluGlySerCysGly 50
Db 906 AGGCGCCGCGGATGACGCGCTGCTGCGACGCGCGCTTCGGA 947
RESULT 26
US-08-222-617A-26/c
; Sequence 26, Application US/08222617A
; Patent No. 5882879
; GENERAL INFORMATION:
; APPLICANT: Veenstra, Annemarie E.
; APPLICANT: Martin, Juan F.
; APPLICANT: Garcia, Bruno D.
; APPLICANT: Gutierrez, Santiago
; APPLICANT: Barredo, Jose L.
; APPLICANT: Von Doehren, Hans
; APPLICANT: Palissa, Harriet
; APPLICANT: Van Liempt, Henk
; APPLICANT: Montenegro, Eduardo P.
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
; NUMBER OF INVENTION: Quantities of ACV Synthetase
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,617A
; FILING DATE: 04-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 97,157
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11444 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-222-617A-26
Alignment Scores:
Pred. No.: 355 Length: 11444
Score: 66.50 Matches: 17
Percent Similarity: 35.44% Conservative: 11
Best Local Similarity: 21.52% Mismatches: 26
Query Match: 20.34% Indels: 25
DB: 2 Gaps: 2
09-833799-13B (1-57) x US-08-222-617A-26 (1-11444)
QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIle----- 19
Db 5877 CAACGAGAAATTTCTGGGCGCCCTTCTCTGCAATCCTTGGCAATCACCACAGACTGTTGA 5818
QY 20 -----LeuIleArgCysAlaMetLeuAsn 27
Db 5817 TTCGTGATAAGACGATAGGATGGCTCAATCTCGCTACTTCGATGCGGAGTCCGCGAA 5758
QY 28 Pro-----ProAsnArgCysLeuLysAspThr 36

Db 5757 TCTTGACCTGGAAGTCATTACGACCTAGATAGTACTCGACCTCCCGCTGCTTCCAGGAATCC 5698
QY 37 AspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 5697 AGCGTACCAGGTCCACCGGTCTTGTTACAAACGGGAGTTAGGACCTTCTCGCTTATCTT 5641
RESULT 27
US-08-222-617A-1/c
; Sequence 1, Application US/08222617A
; Patent No. 5882879
; GENERAL INFORMATION:
; APPLICANT: Veenstra, Annemarie E.
; APPLICANT: Martin, Juan F.
; APPLICANT: Garcia, Bruno D.
; APPLICANT: Gutierrez, Santiago
; APPLICANT: Barredo, Jose L.
; APPLICANT: Von Doehren, Hans
; APPLICANT: Palissa, Harriet
; APPLICANT: Van Liempt, Henk
; APPLICANT: Montenegro, Eduardo P.
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
; NUMBER OF INVENTION: Quantities of ACV Synthetase
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,617A
; FILING DATE: 04-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 97,157
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12364 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Penicillium chrysogenum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 264..11600
; OTHER INFORMATION: /function= "enzyme"
; OTHER INFORMATION: /product= "ACV Synthetase"
US-08-222-617A-1
Alignment Scores:
Pred. No.: 394 Length: 12364
Score: 66.50 Matches: 17
Percent Similarity: 35.44% Conservative: 11
Best Local Similarity: 21.52% Mismatches: 26
Query Match: 20.34% Indels: 25
DB: 2 Gaps: 2
09-833799-13B (1-57) x US-08-222-617A-1 (1-12364)
QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIle----- 19

Db 5877 CAACAGGAATTTCTGGGCCCCCTTCTCTGCAATCCTTGTGCAATCACACAGACTGTTTGA 5818
QY 20 -----LeuileArgCysAlaMetLeuAsn 27
Db 5817 TTCCGTGATAAGACGATAGGCTCAATCTCGCTACTTTCGATGCGCAGTCGCGAA 5758
QY 28 Pro-----ProAsnArgCysLeuLysAspThr 36
Db 5757 TCTTGACCTGGAAGTCATTACGACCTAGACTCGACCTCCCGCTGCTTCAGGAATCC 5698
QY 37 AspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 5697 AGGTACAGGTCACCGGCTTGTACAAACGGGAGGTAGCACCTTCTCGCTTATCTT 5641
RESULT 28
US-07-963-538B-34/c
; Sequence 34, Application US/07963538B
; Patent No. 5951983
; GENERAL INFORMATION:
; APPLICANT: SUGIYAMA, TAKASHI
; APPLICANT: KAMIMURA, TAKASHI
; APPLICANT: MASUDA, KENICHI
; APPLICANT: OKADA, MASAHIRO
; APPLICANT: OHTSUKA, EIKO
; APPLICANT: IMAIZUMI, ATSUSHI
; APPLICANT: WATANABE, KUNIHITO
; APPLICANT: SUGA, TETSUYA
; APPLICANT: MATSUMOTO, YOHICHI
; APPLICANT: TAKEUCHI, AKIKO
; TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
; TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
; TITLE OF INVENTION: TECHNOLOGY
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODDARD LLP
; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/963,538B
; FILING DATE: 20-OCT-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/843,359
; FILING DATE: 25-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,483
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212399
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212398
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-355553
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62-330219
; FILING DATE: 28-DEC-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: TEJN-005/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5070

; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
US-07-963-538B-34
Alignment Scores:
Pred. No.: 0.548 Length: 68
Score: 65.00 Matches: 12
Percent Similarity: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 9
Query Match: 19.88% Indels: 0
DB: 2 Gaps: 0
09-833799-13B (1-57) x US-07-963-538B-34 (1-68)
QY 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 66 ATGCTGAACCCCGCACTTCTGTGAATGGACGGTCAGTGTAAACGAGATCTGAAATGT 7
QY 45 Cys 45
Db 6 TGT 4
RESULT 29
US-08-306-691B-50/c
; Sequence 50, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1364 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-50


```
QY 8 ProValSerThr-----LysProGlySerCysProIleIleLeuIleArg---Cys 23
DB 46094 CCAACCAACAGCGCGATCTTCTCCCGGCC-TGCCCGTACGTGGCGATCAGGCGCTGT 46036

QY 24 AlaMetLeuAsnProProAsn-----ArgCysLeuLys 34
DB 46035 GCCTGATCGGATCACCACCAAGTGTCGCGTCCCATGTGCTCGACCGCATCCACATCG 45976

QY 35 AspThrAspCysProGly-----IleLysLysCysCysGluGlySer 48
DB 45975 GAAACCGACAGCCCGCACTCGCCCAACGCTGGGTGATCACCCGCTGCTGGCAGCGACCA 45916

QY 49 CysGly 50
DB 45915 TTCGGC 45910

RESULT 36
US-09-370-700-1/c
; Sequence 1, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Alignment Scores:
Pred. No.: 1.02e+04 Length: 80161
Score: 64.00 Matches: 20
Percent Similarity: 41.94% Conservative: 6
Best Local Similarity: 32.28% Mismatches: 17
Query Match: 19.57% Indels: 20
DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-370-700-1 (1-80161)

QY 8 ProValSerThr-----LysProGlySerCysProIleIleLeuIleArg---Cys 23
DB 46094 CCAACCAACAGCGCGATCTTCTCCCGGCC-TGCCCGTACGTGGCGATCAGGCGCTGT 46036

QY 24 AlaMetLeuAsnProProAsn-----ArgCysLeuLys 34
DB 46035 GCCTGATCGGATCACCACCAAGTGTCGCGTCCCATGTGCTCGACCGCATCCACATCG 45976

QY 35 AspThrAspCysProGly-----IleLysLysCysCysGluGlySer 48
DB 45975 GAAACCGACAGCCCGCACTCGCCCAACGCTGGGTGATCACCCGCTGCTGGCAGCGACCA 45916

QY 49 CysGly 50
DB 45915 TTCGGC 45910

RESULT 37
US-08-619-198-4/c
; Sequence 4, Application US/08619198
; Patent No. 5885831
```

```
; GENERAL INFORMATION:
; APPLICANT: Young, Michael W.
; APPLICANT: Sehgal, Amita
; APPLICANT: Voshall, Leslie B.
; APPLICANT: Price, Jeffrey L.
; APPLICANT: Myers, Michael
; TITLE OF INVENTION: NUCLEAR LOCALIZATION FACTOR ASSOCIATED
; WITH CIRCADIAN RHYTHMS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,198
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-128A CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4170 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; DESCRIPTION: tim cDNA sequence wherein R at position 1335
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4170
; US-08-619-198-4

Alignment Scores:
Pred. No.: 221 Length: 4170
Score: 63.50 Matches: 14
Percent Similarity: 41.86% Conservative: 4
Best Local Similarity: 32.56% Mismatches: 14
Query Match: 19.42% Indels: 11
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-619-198-4 (1-4170)

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAsp----- 37
DB 3919 AGGTGTATCTCTCCGCTCGCGCGCTGCTCATCAGAGCCACCATGGACAGGAGG 3860

QY 38 -----CysProGlyIleLysLysCysCysGluGlySerCysGly----- 50
DB 3859 AGACGGTGCCCGCGAACCAAGAACTGCTGCTTCTCGTTGTGGCTTCCCATCAAGGTGT 3800

QY 51 MetAlaCys 53
DB 3799 TTAGCATGT 3791

RESULT 38
```


US-08-619-198-8/c
; Sequence 8, Application US/08619198
; Patent No. 5885831
; GENERAL INFORMATION:

APPLICANT: Young, Michael W.
APPLICANT: Sehgal, Amita
APPLICANT: Voshall, Leslie B.
APPLICANT: Price, Jeffrey L.
APPLICANT: Myers, Michael
TITLE OF INVENTION: NUCLEAR LOCALIZATION FACTOR ASSOCIATED
TITLE OF INVENTION: WITH CIRCADIAN RHYTHMS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,198
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-128A CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5192 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

US-08-619-198-8

Alignment Scores:
Pred. No.: 297 Length: 5192
Score: 63.50 Matches: 14
Percent Similarity: 41.86% Conservative: 4
Best Local Similarity: 32.56% Mismatches: 14
Query Match: 19.42% Indels: 11
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-619-198-8 (1-5192)

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAsp----- 37

Db 4398 AGGTGTTACCTCTCCCGCTCCCGCGGTGCTCATCAGAGCCACCATGAGCAGGAGG 4339

QY 38 -----CysProGlyIleLysLysCysCysGluGlySerCysGly----- 50

Db 4338 AGACGGTGCCCGCCAGCAACTGCTGCTGTTCTCGTTGCTTCCCATGAAGGTGT 4279

QY 51 MetAlaCys 53

Db 4278 TTAGCATGT 4270

RESULT 39

US-08-619-198-1/c
; Sequence 1, Application US/08619198
; Patent No. 5885831
; GENERAL INFORMATION:

APPLICANT: Young, Michael W.
APPLICANT: Sehgal, Amita
APPLICANT: Voshall, Leslie B.
APPLICANT: Price, Jeffrey L.
APPLICANT: Myers, Michael
TITLE OF INVENTION: NUCLEAR LOCALIZATION FACTOR ASSOCIATED
TITLE OF INVENTION: WITH CIRCADIAN RHYTHMS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,198
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-128A CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5198 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: tim cDNA sequence wherein R at position 1575
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster

US-08-619-198-1

Alignment Scores:
Pred. No.: 297 Length: 5198
Score: 63.50 Matches: 14
Percent Similarity: 41.86% Conservative: 4
Best Local Similarity: 32.56% Mismatches: 14
Query Match: 19.42% Indels: 11
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-619-198-1 (1-5198)

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAsp----- 37

Db 4396 AGGTGTTACCTCTCCCGCTCCCGCGGTGCTCATCAGAGCCACCATGAGCAGGAGG 4337

QY 38 -----CysProGlyIleLysLysCysCysGluGlySerCysGly----- 50

Db 4336 AGACGGTGCCCGCCAGCAACTGCTGCTGTTCTCGTTGCTTCCCATGAAGGTGT 4277

QY 51 MetAlaCys 53

Db 4276 TTAGCATGT 4268

RESULT 40

US-09-028-934-28
; Sequence 28, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:


```

; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3855 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3855
; OTHER INFORMATION: /note= "nucleic acid sequence with an
; OTHER INFORMATION: open reading frame encoding a delta-182
; OTHER INFORMATION: variant polypeptide"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..2479
; OTHER INFORMATION: /product= "delta-182 variant
; OTHER INFORMATION: polypeptide"
; US-08-974-549A-4
Alignment Scores:
Pred. No.: 267 Length: 3855
Score: 62.50 Matches: 23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match: 19.11% Indels: 25
DB: 4 Gaps: 5
09-833799-13B (1-57) x US-08-974-549A-4 (1-3855)
QY 2 GlncLupProVallysGlyProValSerThrLysProGlySerCysProIleIleLeu 21
Db 1317 CAGCAGCGGTG-----TCTGTGCCGGGAGACAGCCCGCTGTGTGG 1361
QY 22 ArgCysAlaMetLeuAsnProAsnArg----- 31
Db 1362 CGG-----CCCCGAGGAGGAGACAGACCCTGCCTGTGTGCAGC 1406
QY 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
Db 1407 TGCTCCGCCACAGCAGCAGCCCTGTC-----AGGTGTACGGTTCTGTGCGGGCCT 1457
QY 52 AlaCys-----PheValProGln 57
Db 1458 GCCTGCGCGGCTGTGTGCCCCCAG 1481
RESULT 42
US-08-851-843A-224
; Sequence 224, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
```

Db 1407 TGTCCGCGCAGCACAGCAGCCCTGGC-----AGGTGTACGGCTTCGTGCGGGCCT 1457

Qy 52 AlaCys-----PheValProGln 57
||||| :|||
Db 1458 GCCTGCGCGGCTGTGGTCCCGCCAG 1481

RESULT 43
US-08-974-549A-1
; Sequence 1, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4015 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..3454
; OTHER INFORMATION: /product= "hTERT"
; OTHER INFORMATION: /note= "human telomerase reverse
; OTHER INFORMATION: transcriptase (hTERT) catalytic protein
; OTHER INFORMATION: component"
; US-08-974-549A-1

Alignment Scores:
Pred. No.: 282 Length: 4015
Score: 62.50 Matches: 23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match: 19.11% Indels: 25
DB: 4 Gaps: 5

09-833799-13B (1-57) x US-08-974-549A-1 (1-4015)

Qy 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeulle 21
|||:||||| :|||
Db 1317 CAGCAGCGCGTG-----TCTGTCCCGGAGAGAGCCCGGCTGTGTGG 1361

Qy 22 ArgCysAlaMetLeuAsnProProAsnArg----- 31
||| :|||
Db 1362 CGG-----CCCCCGAGGAGGAGACACAGACCCCGTCGCTGTGTGCAGC 1406

Qy 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluCysSerCysGlyMet 51
||| :|||
Db 1407 TGTCGCCGCGCAGCAGCAGCCCTGGC-----AGGTGTACGGCTTCGTGCGGGCCT 1457

Qy 52 AlaCys-----PheValProGln 57
||||| :|||
Db 1458 GCCTGCGCGGCTGTGGTCCCGCCAG 1481

RESULT 44
US-08-854-050-224
; Sequence 224, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:


```

QY 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
Db 1407 TGCTCCGCCAGCACAGCCCTGGC-----AGGTGTACGGCTTCGTGGCGGCT 1457

QY 52 AlaCys-----PheValProGln 57
Db 1458 GCCTGCCCGCGGTGGTGGCCCCAG 1481

RESULT 46
US-09-572-423B-3
; Sequence 3, Application US/09572423B
; Patent No. 6331399
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: William A. Gaarde
; APPLICANT: Edward Wanciewicz
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
; FILE REFERENCE: ISPH-0462
; CURRENT APPLICATION NUMBER: US/09/572,423B
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 3
; LENGTH: 4015
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)...(3454)
US-09-572-423B-3

Alignment Scores:
Pred. No.: 282 Length: 4015
Score: 62.50 Matches: 23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match: 19.11% Indels: 25
DB: 4 Gaps: 5

09-833799-13B (1-57) x US-09-572-423B-3 (1-4015)
QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeulle 21
Db 1317 CAGCAGCCGGTG-----TCTGTGCCCGGAGAGAGCCCGCTCTGTGG 1361

QY 22 ArgCysAlaMetLeuAsnProAsnArg----- 31
Db 1362 CGG-----CCCCGAGGAGGAGACACACACCCCGCTCGCTGGTGCAGC 1406

QY 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
Db 1407 TGCTCCGCCAGCACAGCCCTGGC-----AGGTGTACGGCTTCGTGGCGGCT 1457

QY 52 AlaCys-----PheValProGln 57
Db 1458 GCCTGCCCGCGGTGGTGGCCCCAG 1481

RESULT 47
US-09-128-354-1
; Sequence 1, Application US/09128354
; Patent No. 6337200
; GENERAL INFORMATION:
; APPLICANT: Morin, Gregg B.
; APPLICANT: Geron Corporation
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants
; FILE REFERENCE: 015389-003310US
; CURRENT APPLICATION NUMBER: US/09/128,354
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 08/851,843
; EARLIER FILING DATE: 1997-05-06
; EARLIER APPLICATION NUMBER: US 08/854,050
; EARLIER FILING DATE: 1997-05-09
; EARLIER APPLICATION NUMBER: US 08/911,312

```

```

; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: US 08/912,951
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: US 08/915,503
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: WO PCT/US97/17618
; EARLIER FILING DATE: 1997-10-01
; EARLIER APPLICATION NUMBER: WO PCT/US97/17885
; EARLIER FILING DATE: 1997-10-01
; EARLIER APPLICATION NUMBER: US 08/974,549
; EARLIER FILING DATE: 1997-11-19
; EARLIER APPLICATION NUMBER: US 08/974,584
; EARLIER FILING DATE: 1997-11-19
; EARLIER APPLICATION NUMBER: US 09/052,864
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4015
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)...(3454)
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT) cDNA
US-09-128-354-1

Alignment Scores:
Pred. No.: 282 Length: 4015
Score: 62.50 Matches: 23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match: 19.11% Indels: 25
DB: 4 Gaps: 5

09-833799-13B (1-57) x US-09-128-354-1 (1-4015)
QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeulle 21
Db 1317 CAGCAGCCGGTG-----TCTGTGCCCGGAGAGAGCCCGCTCTGTGG 1361

QY 22 ArgCysAlaMetLeuAsnProAsnArg----- 31
Db 1362 CGG-----CCCCGAGGAGGAGACACACACCCCGCTCGCTGGTGCAGC 1406

QY 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
Db 1407 TGCTCCGCCAGCACAGCCCTGGC-----AGGTGTACGGCTTCGTGGCGGCT 1457

QY 52 AlaCys-----PheValProGln 57
Db 1458 GCCTGCCCGCGGTGGTGGCCCCAG 1481

RESULT 48
US-09-675-321-1
; Sequence 1, Application US/09675321
; Patent No. 6440735
; GENERAL INFORMATION:
; APPLICANT: Gaeta, Federico C.A.
; APPLICANT: Geron Corporation
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune
; TITLE OF INVENTION: Response to a Telomerase Antigen
; FILE REFERENCE: 015389-003500FC
; CURRENT APPLICATION NUMBER: US/09/675,321
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/112,006
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4015

```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(3454)
; OTHER INFORMATION: human telomerase reverse transcriptase (hTRT)
US-09-675-321-1

Alignment Scores:
Pred. No.:      282      Length:      4015
Score:          62.50    Matches:      23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match:      19.11% Indels:      25
DB:               4      Gaps:       5

09-833799-13B (1-57) x US-09-675-321-1 (1-4015)

Qy 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeuLeu 21
Db 1317 CAGCAGCCGGTG-----TCTGTGCCGGGAGAGCCCGCTCTGTGG 1361
Qy 22 ArgCysAlaMetLeuAsnProAsnArg----- 31
Db 1362 CGG-----CCCCGAGGAGGAGACACAGACCCCGCTCGCTGGTGCGAGC 1406
Qy 32 CysLeuLysAspThrAspCysProGlyLysCysCysGluGlySerCysGlyMet 51
Db 1407 TGCTCCGCCAGCAGCAGCCCTGGC-----AGGTGTACGGCTTCGTGGGGCCT 1457
Qy 52 AlaCys-----PheValProGln 57
Db 1458 GCCTGGCCGGCTGTGTGCCCCAG 1481

RESULT 49
US-09-052-919-1
; Sequence 1, Application US/09052919
; Patent No. 644650
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Antisense Compositions for Detecting and
; TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/052,919
; FILING DATE: 31-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
```

```
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/974,549
; FILING DATE: 19-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/974,584
; FILING DATE: 19-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Parent, Annette S.
; REGISTRATION NUMBER: 42,058
; REFERENCE/DOCKET NUMBER: 015389-003600US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4015 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..3454
; OTHER INFORMATION: /product= "human telomerase reverse
; OTHER INFORMATION: transcriptase (hTRT)"
US-09-052-919-1

Alignment Scores:
Pred. No.:      282      Length:      4015
Score:          62.50    Matches:      23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match:      19.11% Indels:      25
DB:               4      Gaps:       5

09-833799-13B (1-57) x US-09-052-919-1 (1-4015)

Qy 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProLleLeuLeu 21
Db 1317 CAGCAGCCGGTG-----TCTGTGCCGGGAGAGCCCGCTCTGTGG 1361
Qy 22 ArgCysAlaMetLeuAsnProAsnArg----- 31
Db 1362 CGG-----CCCCGAGGAGGAGACACAGACCCCGCTCGCTGGTGCGAGC 1406
Qy 32 CysLeuLysAspThrAspCysProGlyLysCysCysGluGlySerCysGlyMet 51
Db 1407 TGCTCCGCCAGCAGCAGCCCTGGC-----AGGTGTACGGCTTCGTGGGGCCT 1457
Qy 52 AlaCys-----PheValProGln 57
Db 1458 GCCTGGCCGGCTGTGTGCCCCAG 1481
```



```
; Sequence 6, Application US/09467997
; Patent No. 6379925
; GENERAL INFORMATION:
; APPLICANT: Kitajewski, Jan
; APPLICANT: Uytendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6677
; TYPE: DNA
; ORGANISM: mouse
US-09-467-997-6

Alignment Scores:
Pred. No.: 561 Length: 6677
Score: 62.50 Matches: 20
Percent Similarity: 40.30% Conservative: 7
Best Local Similarity: 29.85% Mismatches: 11
Query Match: 19.11% Indels: 29
DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-467-997-6 (1-6677)
QY 10 SerThryProGlySerCysProIlelleLeuIleArgCysAlaMetLeuAsn----- 27
Db 3647 TCCCTAAGCCAGGTCACACCGCTC-----TGTCCTGCTCAGTGGTTT 3694
QY 28 -----ProProAsnArgCys 32
Db 3695 GGGGGCCCTGACTGTCTGACACCTCCAGCTCCACCGGCTGGCGTCCCCCTCACCCCTGC 3754
QY 33 LeuLysAsp-----ThrAspCysProGlylleLysLysCysCysgluclySerCys 49
Db 3755 CTCACAATGGTACCTGCTGACCTGAGACCCCTGGTGT-----GGCAACCCG 3799
QY 50 GlyMetAlaCysPheValPro 56
Db 3800 GGCTTTCAATGCACCTGCCCT 3820

RESULT 54
US-08-400-159-3/c
; Sequence 3, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..708
US-08-400-159-3

Alignment Scores:
Pred. No.: 36.4 Length: 708
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 2 Gaps: 1

09-833799-13B (1-57) x US-08-400-159-3 (1-708)
QY 13 ProGlySerCysProIlelleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
Db 175 CCAGGCACCTTGCCG-----TCGTCCTCGTCCGACTCCCCCGCTGC 137
QY 33 LeuLysAspThrAspCysProGlylleLysLysCys 44
Db 136 AGCAGCGACCTGTTGTCCTCCCGCCGTCGTCGTTGC 101

RESULT 55
US-08-611-729A-3/c
; Sequence 3, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..708
US-08-611-729A-3

Alignment Scores:
Pred. No.: 36.4 Length: 708
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 3 Gaps: 1

09-833799-13B (1-57) x US-08-611-729A-3 (1-708)
QY 13 ProGlySerCysProIleLeuIleArgCysAlaMetLeuAnProProAsnArgCys 32
Db 175 CCAGGACATGTCGCG-----TCGCTCCGTCGACACTCCCGCTGC 137
QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 136 ACAGCGACCTCGTTGTCGCCCGCGTATCGTTGC 101

RESULT 56
US-08-152-2/c
; Sequence 2, Application US/08886152
; Patent No. 5880273
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: ARAI, HIROYUKI
; APPLICANT: INOUE, KEIZO
; TITLE OF INVENTION: PLATELET ACTIVATING FACTOR
; TITLE OF INVENTION: ACETYLHYDROLASE, AND GENE THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/886,152
; FILING DATE: 30-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-188369
; FILING DATE: 28-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2292-041-0
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: BOVINE (Bos taurus)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111..1286
US-08-886-152-2

Alignment Scores:
Pred. No.: 116 Length: 1665
Score: 61.50 Matches: 13
Percent Similarity: 42.50% Conservative: 4
Best Local Similarity: 32.50% Mismatches: 14
Query Match: 18.81% Indels: 9
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-886-152-2 (1-1665)
QY 13 ProGlySerCysProIleLeuIleArgCysAlaMetLeu----- 26
Db 823 CCAAAATGAATGCCATTACAGCCACACGCTCAGTCAATGCCGCTTCAAGTCAATC 764
QY 27 -----AsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 763 AGATCCAATCCGCCAGCAAGATGTTGAGA---ACGGCTGCCACGACGTGACCTCTTGT 707

RESULT 57
US-09-196-222-2/c
; Sequence 2, Application US/09196222
; Patent No. 6323017
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: ARAI, HIROYUKI
; APPLICANT: INOUE, KEIZO
; TITLE OF INVENTION: PLATELET ACTIVATING FACTOR
; TITLE OF INVENTION: ACETYLHYDROLASE, AND GENE THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/886,152
; FILING DATE: 30-JUN-1997
; APPLICATION NUMBER: JP 8-188369
; FILING DATE: 28-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2292-041-0
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: BOVINE (Bos taurus)
FEATURE:
NAME/KEY: CDS
LOCATION: 111..1286
US-09-196-222-2

Alignment Scores:
Pred. No.: 116 Length: 1665
Score: 61.50 Matches: 13
Percent Similarity: 42.50% Conservative: 4
Best Local Similarity: 32.50% Mismatches: 14
Query Match: 18.81% Indels: 9
DB: 4 Gaps: 2

09-833799-13B (1-57) x US-09-196-222-2 (1-1665)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeu----- 26
Db 823 CCAATGAATGTCCTTACAGCCACAGCGCTCACGTCAATGCGCGCTTCAAGGTATC 764
QY 27 -----AsnProProAsnArgCysLeuIleAspThrAspCysProGlyIleLysCys 44
Db 763 AGATCCAAATCGCCAGCAAGATGTTGAGA---ACGCGCTGCCAGCAGTGACCTCTTGT 707

RESULT 58

US-08-264-534-5/c
Sequence 5, Application US/08264534
Patent No. 5648464
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon et al.
TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,534
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/695,189
FILING DATE: 03-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 2892 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2640
US-08-264-534-5

Alignment Scores:

Pred. No.: 244 Length: 2892
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 1 Gaps: 1

09-833799-13B (1-57) x US-08-264-534-5 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
Db 316 CCAGGCACTTGCCG-----TCGCTCCGTCGCGACTCCCCGCTGC 278
QY 33 LeuIleAspThrAspCysProGlyIleLysLysCys 44
Db 277 AGCAGCGACCTCGTGTGTCGCCGCGGTGATCGTTGC 242

RESULT 59

US-08-083-590A-1/c
Sequence 1, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2892 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2640
US-08-083-590A-1

Alignment Scores: 244 Length: 2892
 Pred. No.: 61.50 Matches: 12
 Score: 43.75% Conservative: 2
 Percent Similarity: 37.50% Mismatches: 11
 Best Local Similarity: 37.50% Indels: 7
 Query Match: 18.81% Gaps: 1
 DB: 1

09-833799-13B (1-57) x US-08-083-590A-1 (1-2892)

Qy 13 ProGlySerCysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCys 32
 |||||:|||||
 Db 316 CCAGGCACTTGCCTGCTCCGCTCCGACTCCCGCTGC 278
 |||||:|||||

Qy 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
 |||||:|||||
 Db 277 AGCAGGACCTCGTGTGCTCCCGCTGATCGTTGC 242
 |||||:|||||

RESULT 60
 US-08-465-500-5/c
 ; Sequence 5, Application US/08465500
 ; Patent No. 5789195
 ; GENERAL INFORMATION:
 ; APPLICANT: Artavanis-Tsakonas, Spyridon
 ; APPLICANT: Muskavitch, Marc A.T.
 ; APPLICANT: Fehon, Richard G.
 ; APPLICANT: Rebay, Ilaria
 ; APPLICANT: Blauweller, Christine M.
 ; APPLICANT: Shepard, Scott B.
 ; TITLE OF INVENTION: HUMAN NOTCH AND DELTA, BINDING DOMAINS
 ; TITLE OF INVENTION: IN TOPOTHYMIC PROTEINS, AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/465,500
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-034
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2892 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 142..2640
 ; US-08-465-500-5

Alignment Scores: 244 Length: 2892
 Pred. No.: 61.50 Matches: 12
 Score: 43.75% Conservative: 2
 Percent Similarity: 37.50% Mismatches: 11
 Best Local Similarity: 37.50%

Query Match: 18.81% Indels: 7
 DB: 1 Gaps: 1

09-833799-13B (1-57) x US-08-465-500-5 (1-2892)

Qy 13 ProGlySerCysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCys 32
 |||||:|||||
 Db 316 CCAGGCACTTGCCTGCTCCGCTCCGACTCCCGCTGC 278
 |||||:|||||

Qy 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
 |||||:|||||
 Db 277 AGCAGGACCTCGTGTGCTCCCGCTGATCGTTGC 242
 |||||:|||||

RESULT 61
 US-08-346-126-5/c
 ; Sequence 5, Application US/08346126
 ; Patent No. 5849869
 ; GENERAL INFORMATION:
 ; APPLICANT: Artavanis-Tsakonas, Spyridon et al.
 ; TITLE OF INVENTION: Human No. 5849869ch And Delta, Binding Domains
 ; TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/346,126
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/791,923
 ; FILING DATE: 14-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-007
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 8698864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2892 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 142..2640
 ; US-08-346-126-5

Alignment Scores: 244 Length: 2892
 Pred. No.: 61.50 Matches: 12
 Score: 43.75% Conservative: 2
 Best Local Similarity: 37.50% Mismatches: 11
 Query Match: 18.81% Indels: 7
 DB: 2 Gaps: 1

09-833799-13B (1-57) x US-08-346-126-5 (1-2892)

Qy 13 ProGlySerCysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCys 32
 |||||:|||||

```

US-08-532-384-1/c
; Sequence 1, Application US/08532384
; Patent No. 6083904
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,384
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,590
; FILING DATE: 25-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leelie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2640
US-08-532-384-1

Alignment Scores:
Pred. No.: 244 Length: 2892
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 3 Gaps: 1

09-833799-13B (1-57) x US-08-532-384-1 (1-2892)

QY 13 ProGlyserCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
Db 316 CCAGGACACTTGCCG-----TGCTCCGTCGCACTCCCGCTGC 278
QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 277 AGCAGCGACCCCTGTTGTCCCGCGTGATCGTTGC 242

RESULT 64
US-08-893-828-5/c
; Sequence 5, Application US/08893828
; Patent No. 6090922
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Muskavitch, Marc A.T.

```


STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSK/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
US-08-728-323A-4

Alignment Scores:
Pred. No.: 36.2 Length: 631
Score: 61.00 Matches: 18
Percent Similarity: 45.61% Conservative: 8
Best Local Similarity: 31.58% Mismatches: 22
Query Match: 18.65% Indels: 9
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-728-323A-4 (1-631)

QY 1 AlaGlnGluProValIysGlyProValSerThrIysProGlySerCysProIleLeu 20
Db 475 GCACGGGGTCCCGCGCGGTCGCGTGCATACAGCGCG----- 513
QY 21 IleArgCysAlaMetLeuAsnProPro---AsnArg-CysLeuIysAspThrAspCysPr 39
Db 514 GTGAGGGCCCCCTGGTCTGTCGCCCTCGAAACACAGGGTGTGTGAAACACAGGTTGCCA 573
QY 39 oGlyIleLysIysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 574 AGCGCGCAATACCCCTCTGCACGGTGTGTGGACGTGGGTGTATGCTC 622

RESULT 67

US-08-998-416-149/c
Sequence 149, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Rebschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBVA GOSSYP11
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park

STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 149:
SEQUENCE CHARACTERISTICS:
LENGTH: 653 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: PAG1057RP
ORGANISM: US-08-998-416-149

Alignment Scores:
Pred. No.: 37.9 Length: 653
Score: 61.00 Matches: 11
Percent Similarity: 63.64% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 8
Query Match: 18.65% Indels: 0
DB: 4 Gaps: 0

09-833799-13B (1-57) x US-08-998-416-149 (1-653)

QY 7 GlyProValSerThrIysProGlySerCysProIleLeuIleArgCysAlaMetLeu 26
Db 275 GGACCGGTGCCACCCACCGTCTCCTCCCTCGCTGTTGCTCGCTGCTGCTA 216
QY 27 AsnPro 28
Db 215 AACCCA 210

RESULT 68

US-09-056-556-191
Sequence 191, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30


```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-6031
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 191:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-056-556-191

Alignment Scores:
Pred. No.: 72.4 Length: 1055
Score: 61.00 Matches: 20
Percent Similarity: 32.10% Conservative: 6
Best Local Similarity: 24.69% Mismatches: 19
Query Match: 18.65% Indels: 36
DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-056-556-191 (1-1055)

QY 3 GluProValLysGlyProValSer-----ThrLysProGlySer 15
Db 275 AAACCCGCGCCAGCGGCGCTGACGGGCTCATGGCAGCGAAATTAGAAACCCGGGATAT 334
QY 16 CysProIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu----- 33
Db 335 TGTCGCGGATGTGCATACGATGCTGAGTGTCTGT-----GGTTCGTGTTTAGCCATT 388
QY 34 -----LysAspThrAspCys----- 38
Db 389 GAGTGTGGATGTTGAGACCCCTGGCCTGGAAGGGGACAAACGCTGCTTTTGCTCTTGGTC 448
QY 39 -----ProGlyIleLysLysCysCysGluGlySer 48
Db 449 CGCCTTTGCGCCGACGCGGTGGTGGCGAAACGCGCTGAGTCGGGAATGCTCGCGCGGT 508
QY 49 Cys 49
Db 509 TGT 511

RESULT 69
US-09-072-596-186
; Sequence 186, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
;

;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-072-596-186

Alignment Scores:
Pred. No.: 72.4 Length: 1055
Score: 61.00 Matches: 20
Percent Similarity: 32.10% Conservative: 6
Best Local Similarity: 24.69% Mismatches: 19
Query Match: 18.65% Indels: 36
DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-072-596-186 (1-1055)

QY 3 GluProValLysGlyProValSer-----ThrLysProGlySer 15
Db 275 AAACCCGCGCCAGCGGCGCTGACGGGCTCATGGCAGCGAAATTAGAAACCCGGGATAT 334
QY 16 CysProIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu----- 33
Db 335 TGTCGCGGATGTGCATACGATGCTGAGTGTCTGT-----GGTTCGTGTTTAGCCATT 388
QY 34 -----LysAspThrAspCys----- 38
Db 389 GAGTGTGGATGTTGAGACCCCTGGCCTGGAAGGGGACAAACGCTGCTTTTGCTCTTGGTC 448
QY 39 -----ProGlyIleLysLysCysCysGluGlySer 48
Db 449 CGCCTTTGCGCCGACGCGGTGGTGGCGAAACGCGCTGAGTCGGGAATGCTCGCGCGGT 508
QY 49 Cys 49
Db 509 TGT 511

RESULT 70
US-09-615-192A-172/c
; Sequence 172, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 1000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
;

```

```

; SEQ ID NO 172
; LENGTH: 518
; TYPE: DNA
; ORGANISM: P
US-09-615-192A-

```

Alignment Scores:

Pred. No.:	37.3	Length:	518
Score:	60.00	Matches:	17
Percent Similarity:	57.89%	Conservative:	5
Best Local Similarity:	44.74%	Mismatches:	14
Query Match:	18.35%	Indels:	3
DB:	4	Gaps:	1

09-833799-13B (1-57) x US-09-615-192A-172 (1-518)

QY	13	ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys	32
Db	211	CTGTAGCTGCCATTGACGGTGATTAT-TGTTATTCCTTGCATACGCTGGTAACGT	153
QY	33	LeuIysAspThrAspCysProGlyIleLysLysCys-----CysGluGlySer	48
Db	152	CTTCGCTTCATGATGAAGGTATAATTGACGATTGCTGCATTGTGATGGCAGC	99

RESULT 71

US-09-449-285A-8
; Sequence 8, Application US/09449285A
; Patent No. 6313280

GENERAL INFORMATION:

APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnologie
 TITLE OF INVENTION: SMAD-INTERACTING POLYPEPTIDES AND THEIR USE
 FILE REFERENCE: 2676-4232US
 CURRENT APPLICATION NUMBER: US/09/449,285A
 CURRENT FILING DATE: 1999-11-24
 PRIOR APPLICATION NUMBER: PCT/EP98/03193
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 97201645.5
 PRIOR FILING DATE: 1997-06-02
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 8
 LENGTH: 850
 TYPE: DNA
 ORGANISM: Mus musculus
 S-09-449-285A-8

Alignment Scores:

Pred. No.:	72.7	Length:	850
Score:	60.00	Matches:	19
Percent Similarity:	39.06%	Conservative:	6
Best Local Similarity:	29.69%	Mismatches:	23
Query Match:	18.3%	Indels:	16
DB:	4	Gaps:	4

09-833799-13B (1-57) x US-09-449-285A-8 (1-850)

4	ProValIysGlyProValSerThrIysProGlySerCysProIleLeu	-----Leu 20
65	CCCCCTCCGGCCCTGCCCGAACTCCAGCCCGAGCGCTGTACTGCCAGATACA	124
21	IleArgCysAlaMet	-----LeuAsnProProAsnArgCys 32
125	GCAAGATGCCGATCTCTGGCAGCAGACACGGCGAGCACTGTCCCCGGTCCCGCAGCC	184
33	LeuIysAspThrAspCysProGlyIleIysIysCysCysGluGlySerCysGlyMetAla	52
185	CTG-----GCCCTAGCCCCCAGCGCTGCTGCCCTGCATCA---	GGGAGGGCC 229
53	CysPheValPro	56
230	GCGGAGACCCCA	241

RESULT 72

```

US-08-858-207A-161/c
; Sequence 161, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858, 207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IS-08-858-207A-161

```

Alignment Scores:	
Pred. No.:	102
Scores:	60.00
Percent Similarity:	43.75%
Best Local Similarity:	27.08%
Query Match:	18.35%
DB:	4
Length:	1089
Matches:	13
Conservative:	8
Mismatches:	11
Indels:	16
Gaps:	1

09-833799-13B (1-57) x US-08-858-207A-161 (1-1089)

Qy	8	ProValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn	27
Db	912	CCAAAGCTTCTGCCAGATGGTCTTGCCCAATAATTCTAATGAANAATCGSAGTGACCATG	853
Qy	28	ProProAsnArgCysLeu-----	33
Db	852	GCgcctgaaaagcactcaatgtgtcagacagaggTTTAAGCAGACAATAAAGGTtACT	793
Qy	34	-----LysAspThrAspCysPro	39
Db	792	AGCAAACAGGAAAAAGACATGccCT	769

RESULT 73
US-09-282-147-37/c
; Sequence 37, Application US/09282147
; Patent No. 6274147
; GENERAL INFORMATION:

APPLICANT: VAKHARIA, Vikram
APPLICANT: YAO, Kun
TITLE OF INVENTION: METHOD FOR GENERATING NONPATHOGENIC, INFECTIOUS
TITLE OF INVENTION: PANCREATIC NECROSIS VIRUS (IPNV) FROM SYNTHETIC RNA
TITLE OF INVENTION: TRANSCRIPTS
FILE REFERENCE: 8288-9023
CURRENT APPLICATION NUMBER: US/09/282,147
EARLIER FILING DATE: 1999-03-31
EARLIER APPLICATION NUMBER: US/60/080,278
EARLIER FILING DATE: 1998-03-31
EARLIER APPLICATION NUMBER: PCT/US97/12955
EARLIER FILING DATE: 1998-03-31
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 2783
TYPE: DNA
ORGANISM: Infectious pancreatic necrosis virus
US-09-282-147-37

Alignment Scores: 360 Length: 2783
Pred. No.: 60.00 Matches: 15
Score: 41.67% Conservative: 5
Best Local Similarity: 31.25% Mismatches: 26
Query Match: 18.35% Indels: 2
DB: 4 Gaps: 1

09-833799-13B (1-57) x US-09-282-147-37 (1-2783)

QY 8 ProValSerThrIysProGlySerCysProIleLeuLeuArgCysAlaMetLeuAsn 27
Db 1376 CCTGGTCAGAGGTAGTACATCATGCGCTGCATGTTGTGGAGTGCAGTTGGCCTCGC 1317
QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysCysCys----- 45
Db 1316 CTTTTCCTAGTCTATTAGTACACCGTGTGTCCTGCAGTATGTAGATGTTGTCTGCAT 1257
QY 46 GluGlySerCysGlyMetAlaCys 53
Db 1256 AGATCATCATCAGTGTCTCTGTTGT 1233

RESULT 74
US-09-543-084A-29/c
Sequence 29, Application US/09543084A
Patent No. 6361988
GENERAL INFORMATION:
APPLICANT: Frances H. Arnold
APPLICANT: Zhixin Shao
APPLICANT: Huimin Zhao
APPLICANT: Lorraine J. Giver
TITLE OF INVENTION: Recombination of Polynucleotide
TITLE OF INVENTION: Sequences Using Defined or Random Primer Sequences
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 98
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/543,084A
FILING DATE: April 4, 2000
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/381,935
APPLICATION NUMBER: 60/041,666

FILING DATE: March 25, 1997
APPLICATION NUMBER: 60/045,211
FILING DATE: April 30, 1997
APPLICATION NUMBER: 60/046,256
FILING DATE: May 12, 1997
APPLICATION NUMBER: 08/905,359
FILING DATE: August 4, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 330187-89
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 788-5100
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 3222 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polynucleotide
US-09-543-084A-29

Alignment Scores: 439 Length: 3222
Pred. No.: 60.00 Matches: 15
Score: 40.00% Conservative: 5
Best Local Similarity: 30.00% Mismatches: 16
Query Match: 18.35% Indels: 14
DB: 4 Gaps: 3

09-833799-13B (1-57) x US-09-543-084A-29 (1-3222)

QY 16 CysProIleLeuLeuArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35
Db 1209 TGCCCGATC-----TCGATGATCGGTGCGCCGACGACGCGCGCCCTCG 1165
QY 36 ThrAsp---CysProGlyIleLysCys-----CysGlu 46
Db 1164 ACGTCGTAGCGCGCGGCACCTTGAGTGCATCCGTTAGAGCGTGGCGCGCCCTGCCAC 1105
QY 47 GlySerCysGlyMetAlaCysPheValPro 56
Db 1104 GGGAAAGTGGGGTTGCCCGACGACCATCCCG 1075

RESULT 75
US-09-543-084A-26/c
Sequence 26, Application US/09543084A
Patent No. 6361988
GENERAL INFORMATION:
APPLICANT: Frances H. Arnold
APPLICANT: Zhixin Shao
APPLICANT: Huimin Zhao
APPLICANT: Lorraine J. Giver
TITLE OF INVENTION: Recombination of Polynucleotide
TITLE OF INVENTION: Sequences Using Defined or Random Primer Sequences
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 98
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/543,084A
FILING DATE: April 4, 2000
CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/381,935
;; APPLICATION NUMBER: 60/041,666
;; FILING DATE: March 25, 1997
;; APPLICATION NUMBER: 60/045,211
;; FILING DATE: April 30, 1997
;; APPLICATION NUMBER: 60/046,256
;; FILING DATE: May 12, 1997
;; APPLICATION NUMBER: 08/905,359
;; FILING DATE: August 4, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oldenkamp, David J.
;; REGISTRATION NUMBER: 29,421
;; REFERENCE/DOCKET NUMBER: 330187-89
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (310) 788-5000
;; TELEFAX: (310) 788-5100
;; INFORMATION FOR SEQ ID NO: 26:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4050 nucleotides
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: polynucleotide
US-09-543-084A-26

Alignment Scores:
Pred. No.: 598 Length: 4050
Score: 60.00 Matches: 15
Percent Similarity: 40.00% Conservative: 5
Best Local Similarity: 30.00% Mismatches: 16
Query Match: 18.35% Indels: 14
DB: 4 Gaps: 3

09-833799-13B (1-57) x US-09-543-084A-26 (1-4050)
Qy 16 CysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35
Db 2083 TCGCCGATC-----TCGATGATCGGGTCGCCGACCGCGCGCCCTCG 2039
Qy 36 ThrAsp---CysProGlyIleLysLysCys-----CysGlu 46
Db 2038 ACGTCGTAGCGCGCGGCACCTTGAGGTGCATCCGGTAGAGCGTGGCGCGCCCTGCCAC 1979
Qy 47 GlySerCysGlyMetAlaCysPheValPro 56
Db 1978 GGGAAAGTGGGGTTGGCCAGCACCATCCCG 1949

Search completed: February 15, 2003, 21:14:41
Job time : 101 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 21:53:32 ; Search time 1184 Seconds
(without alignments)
2817.796 Million cell updates/sec

Title: 09-833799-13A

Perfect score: 206

Sequence: 1 AATTCGACTCGGTACCAT.....GTTTCGTTCCACAATAATAG 206

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gss.*
- 18: em_gss_hum.*
- 19: em_gss_inv.*
- 20: em_gss_pln.*
- 21: em_gss_vrt.*
- 22: em_gss_fun.*
- 23: em_gss_nam.*
- 24: em_gss_mus.*
- 25: em_gss_other.*
- 26: em_gss_pro.*
- 27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	126.2	61.3	388	12	BE772159	
C	2	126.2	61.3	423	9	AT830872
C	3	126.2	61.3	438	10	AW001880
C	4	124.6	60.5	219	10	BE182652
C	5	124.6	60.5	340	13	EG987761
C	6	124.6	60.5	342	10	BE182993
						BE772159 CM4-FT010
						AT830872 wj61c08.x
						AW001880 wj61c08.x
						BE182652 RC3-HT064
						EG987761 CM3-HT119
						BE182993 CM4-HT065

BE181546	CM0-HT063
BE182864	CM4-HT065
BE772161	CM4-FT010
BE715602	CM4-HT074
BE715605	CM4-HT074
BE715626	CM4-HT074
BE772160	CM4-FT010
BF838555	IL5-HT099
BE772166	CM4-FT010
AI285400	ql69f06.x
BE715603	CM4-HT074
BE772155	CM4-FT010
BE772171	CM4-FT010
BE182879	CM4-HT065
BE772169	CM4-FT010
AA583997	nm65801.s
BE772168	CM4-FT010
AI33896	tg05609.x
AI989737	wu21608.x
BE715611	CM4-HT074
BE715615	CM4-HT074
BE772137	CM4-FT010
AA586974	nm69808.s
BE715592	CM4-HT074
BE715609	CM4-HT074
AI874186	wm50802.x
AW137392	UI-H-B11-
BE715604	CM4-HT074
AA586943	nm69801.s
AW082097	xb60d08.x
BE772154	CM4-FT010
AI283910	qt66a10.x
BF836224	QV3-HT101
BE715601	CM4-HT074
BF836743	CM2-HT096
AI582329	tg66909.x
AI459240	tk11c09.x
BF824934	IL0-HN003
BF825207	IL0-HN003
BF837734	QV3-HT101
BF824932	IL0-HN003
AA582866	nm72809.s
BF002099	7g99b04.x
BF837735	QV3-HT101
BQ941085	AGENCOURT
BG490472	602519610
AI392753	tg23903.x
AI858070	wj70803.x
AI807596	we43c11.x
BE711100	RC3-HT064
BE181373	CM0-HT063
BE181959	CM1-HT064
BE715619	CM4-HT074
BF094327	CM4-UT004
BF836201	QV3-HT101
BI021805	CM3-MT031
AI924155	wn64c08.x
BI061069	IL3-UT011
AI830232	wj78912.x
BF835679	QV3-HT101
AW845140	IL0-CT000
AW845142	IL0-CT000
BF824930	IL0-HN003
BF824935	IL0-HN003
BF824920	IL0-HN003
BF825204	IL0-HN003
AA5860590	QV0-CT038
AA586983	nm69810.s
BE736734	PM1-KT000
BM768573	K2-EST051
BF836725	CM2-HT096
BF825196	IL0-HN003
BE711083	RC3-HT064

80 120.8 58.6 344 10 AW844969 MR0-CT000
 C 81 120.2 58.3 391 12 BE772162 CM4-FT010
 82 120 58.3 298 12 BF937690 QV3-HT101
 83 120 58.3 313 12 BF736294 QV3-HT001
 84 120 58.3 314 12 BF837677 QV3-HT101
 C 85 119.8 58.2 431 12 BF837730 QV3-HT101
 86 116.6 56.6 470 12 BF913875 MR3-UT010
 C 87 115.4 56.0 448 10 AW082007 xb59h01.x
 C 88 114.6 55.6 289 12 BF934601 CM2-HT096
 C 89 113.6 55.1 275 10 BE182850 RC3-HT064
 C 90 113.6 55.1 275 12 BE711090 RC3-HT064
 91 113.6 55.1 393 10 BE183902 CM0-HT067
 92 113.6 55.1 393 12 BE772164 CM4-FT010
 C 93 112.6 54.7 400 9 AI831796 wj40h05.x
 C 94 112.6 54.7 411 9 AA586718 nm67b11.s
 C 95 112.6 54.7 470 12 BF913869 MR3-UT010
 C 96 112.6 54.7 478 9 AA527557 ng42a03.s
 C 97 112.6 54.6 555 9 AA583567 nm61b01.s
 98 112.2 54.5 468 12 BF090376 QV3-NT002
 99 111.8 54.3 336 12 BF844574 RC5-HT098
 C 100 110 53.4 190 10 BE181433 CM0-HT063

ALIGNMENTS

RESULT 1
 BE772159 388 bp mRNA linear EST 20-SEP-2000
 LOCUS CM4-FT0104-230600-215-c06 FT0104 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE772159
 ACCESSION BE772159
 VERSION BE772159.1 GI:10225817
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 388)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., Geolivera, P.S., Sucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, B.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=CM4-FT0104-230
 600-215-c06&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 26
 High quality sequence stop: 388.
 Location/Qualifiers

FEATURES
 source
 1..388
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORSTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 88 a 115 c 93 g 92 t
 ORIGIN

Query Match 61.3%; Score 126.2; DB 12; Length 388;
 Best Local Similarity 81.6%; Pred. No. 8.8e-30;
 Matches 146; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
 QY 28 ATGCTCAGAACCCAGTAAAGGTCTGTGCTACTAGCCAGGTTCTGCTATTATCT 87
 Db 113 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCACTAAGCCCTGCTGCTATTATCT 172
 QY 88 TGATTCTGTTGGCTATGTTAAACCCACCTAACCCTTGTTTGAAGGACACTGATTGTCCAG 147
 Db 173 TGATCCGTTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATATCTGACTGCCAG 232
 QY 148 GTATCAAAAGTGTGTGAAGGTTCTCTCGGTATGGCTTTCCTCCCAATAATAG 206
 Db 233 GAATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTTCCTCCCAAGTGAGAG 291

RESULT 2
 AI830872/c
 LOCUS AI830872 423 bp mRNA linear EST 21-DEC-1999
 DEFINITION wj61c08.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2407310 3,
 similar to gb:S58717 tnal ELAFIN PRECURSOR (HUMAN); contains element
 MER22 repetitive element ; mRNA sequence.

ACCESSION AI830872
 VERSION AI830872.1 GI:5451543
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 423)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennox, Ph.D.
 Cloning: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.lnl.gov/bbrp/image/image.html
 Insert Length: 534 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES
 source
 1..423
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2407310"
 /clone_lib="NCI CGAP Lul9"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DHI0B (phage-resistant)"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

```

BASE COUNT      106 a      91 c      125 g      101 t
ORIGIN
Query Match      61.3%; Score 126.2; DB 9; Length 423;
Best Local Similarity 81.6%; Pred. No. 9.1e-30;
Matches 146; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 361 AAGCGCAAGAGCCAGTCAAAAGGTCCTCAATAGCCCTGCTCTGCCCATTTATCT 302

QY 88 TGATTCCTGGCTATCTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 301 TGATCCGTCGCCAATGTTGAATCCCTAACCGTTGTTTGAAGAACTGACTGCCAG 242

QY 148 GTATCAAAAAGTGTGTGAAGGTTCTTCCGGTATGGCTGTTTTCGTTCCACATAATAG 206
Db 241 GAATCAAGAAGTGTGTGAAGGTTCTTCCGGGATGGCTGTTTTCGTTCCCGAGTGAAG 183

RESULT 3
AW001880/c
LOCUS
DEFINITION
IMAGE:2513867.3' Similar to gb:S58717, rnal ELAFIN PRECURSOR (HUMAN)
);contains element MER28 repetitive element ;, mRNA sequence.
ACCESSION
AW001880
VERSION
AW001880.1 GI:5848796
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 438)
AUTHORS
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 552, Std Error: 0.00
Seq primer: -400P from Gibco.
FEATURES
Location/Qualifiers
1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2513867"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: thymus, pooled; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15,
TGTTCACCAATCTGAAGTCGGAGCGCGGCAAGCTTTTTTTTTTTT 3' ],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      106 a      92 c      125 g      115 t
ORIGIN
Query Match      61.3%; Score 126.2; DB 10; Length 438;
Best Local Similarity 81.6%; Pred. No. 9.2e-30;
Matches 146; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 377 AAGCGCAAGAGCCAGTCAAAAGGTCCTGTTTCCACTAAGCTGGCTCTGCCCATTTATT 318

```

```

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 317 TGATCCGTCGCCAATGTTGAATCCCTAACCGCTGCTTGAAGATACTGACTGCCAG 258

QY 148 GTATCAAAAAGTGTGTGAAGGTTCTTCCGGTATGGCTGTTTTCGTTCCACATAATAG 206
Db 257 GAATCAAAAAGTGTGTGAAGGTTCTTCCGGGATGGCTGTTTTCGTTCCCGAGTGAAG 199

RESULT 4
BE182652
LOCUS
DEFINITION
RC3-HU0649-090500-011-h03 HU0649 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE182652
VERSION
BE182652.1 GI:8661828
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 219)
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=RC3-HU0649-090
500-011-h03&t3=2000-05-09&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 219
High quality sequence stop: 219.
FEATURES
Location/Qualifiers
1..219
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/notes="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      51 a      60 c      59 g      49 t
ORIGIN
Query Match      60.5%; Score 124.6; DB 10; Length 219;
Best Local Similarity 81.0%; Pred. No. 2.4e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 22 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCCTGGCTCTGCCCATTTATCT 81

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 82 TGATCCGTCGCCAATGTTGAATCCCTAACCGCTGCTTGAAGATACTGACTGCCAG 141

```

QY 148 GTATCAAAAGTCTGTGAAGGTTCTTCGGGTATGGCTTCTTCGTTCCACAAATATAG 206
 Db 142 GAATCAAGAAGTCTGTGAAGGCTTTCGGGATGGCTTCTTCGTTCCCAAGTGAGAG 200

RESULT 5
 BG987761/c 340 bp mRNA linear EST 13-JUN-2001
 LOCUS CM3-HT1190-110101-624-h04 HT1190 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION BG987761
 VERSION BG987761.1 GI:14391831
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 340)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM3&t2=CM3-HT1190-110101-624-h04&t3=2001-01-11&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 6
 High quality sequence stop: 340.
 FEATURES
 source
 1..340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT1190"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 79 a 82 c 97 g 82 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 13; Length 340;
 Best Local Similarity 81.0%; Pred. No. 2.8e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTTATATCT 87
 Db 240 AAGCGCAAGACCAAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATATCT 181
 QY 88 TGATTCTGGCTATGTTAAACCACTAACCGTTGTTTGAAGGACACTGATGTCAG 147
 Db 180 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGAC 121
 QY 148 GTATCAAAAGTCTGTGAAGGTTCTTCGGGTATGGCTTCTTCGTTCCACAAATATAG 206

Db 120 GAATCAAGAAGTCTGTGAAGGCTTTCGGGATGGCTTCTTCGTTCCCAAGTGAGAG 62

RESULT 6
 BE182993/c 342 bp mRNA linear EST 22-JUN-2000
 LOCUS CM4-HT0653-180400-146-a11 HT0653 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION BE182993
 VERSION BE182993.1 GI:8662169
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 342)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM4-HT0653-180400-146-a11&t3=2000-04-18&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 342.
 FEATURES
 source
 1..342
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0653"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 80 a 82 c 100 g 79 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 10; Length 342;
 Best Local Similarity 81.0%; Pred. No. 2.8e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTTATATCT 87
 Db 261 AAGCGCAAGACCAAGTCAAGGTCAGTCTCCACTAAGCCCTGGCTCTGCCCATATCT 202
 QY 88 TGATTCTGGCTATGTTAAACCACTAACCGTTGTTTGAAGGACACTGATGTCAG 147
 Db 201 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGAC 142
 QY 148 GTATCAAAAGTCTGTGAAGGTTCTTCGGGTATGGCTTCTTCGTTCCACAAATATAG 206
 Db 141 GAATCAAAAGTCTGTGAAGGCTTTCGGGATGGCTTCTTCGTTCCCAAGTGAGAG 83

RESULT 7
 BE181546/c
 LOCUS
 DEFINITION CM0-HT0634-270300-308-f02 HT0634 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE181546
 VERSION BE181546.1 GI:86607722
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 344)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0634-270300-308-f02&t3=2000-03-27&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 6
 High quality sequence stop: 344.
 Location/Qualifiers
 1..344
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0634"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 79 a 80 c 113 g 72 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 10; Length 344;
 Best Local Similarity 81.0%; Pred. No. 2.8e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTATTATCT 87
 DB 296 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCT 237
 QY 88 TGATTGGTGGCCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
 DB 236 TGATCGGTGGCCCATGTTGAATCCCTTAAACCGTGTCTTGAAGATACTGACTGCCAG 177
 QY 148 GTATCAAAAAGTGTGAAGGTTCTCGGTATGCGTTGTTTCCTCCACAATAATAG 206
 DB 176 GAATCAAGAAGTGTGAAGGCTCTTGGGATGGCCTGTTTCTGCCAGTGAGAG 118
 RESULT 8
 BE182864
 LOCUS
 DEFINITION CM4-HT0652-150400-143-a07 HT0652 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE182864
 VERSION BE182864.1 GI:8662040
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 359)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0652-150400-143-a07&t3=2000-04-15&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 39
 High quality sequence stop: 359.
 Location/Qualifiers
 1..359
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0652"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 92 a 90 c 95 g 82 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 10; Length 359;
 Best Local Similarity 81.0%; Pred. No. 2.8e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTATTATCT 87
 DB 170 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCT 229
 QY 88 TGATTGGTGGCCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
 DB 230 TGATCGGTGGCCCATGTTGAATCCCTTAAACCGTGTCTTGAAGATACTGACTGCCAG 289
 QY 148 GTATCAAAAAGTGTGAAGGTTCTCGGTATGCGTTGTTTCCTCCACAATAATAG 206
 DB 290 GAATCAAGAAGTGTGAAGGCTCTTGGGATGGCCTGTTTCTGCCAGTGAGAG 348
 RESULT 9
 BE772161/c
 LOCUS
 DEFINITION CM4-FT0104-230600-215-c09 FT0104 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE772161
 VERSION BE772161.1 GI:10225819

```

KEYWORDS      EST.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 361)
               Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
               Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
               Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
               Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
               , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
               Simpson, A.J.
TITLE          Shotgun sequencing of the human transcriptome with ORF expressed
               sequence tags
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE        20202663
COMMENT        Contact: Simpson A.J.G.
               Laboratory of Cancer Genetics
               Ludwig Institute for Cancer Research
               Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
               Brazil
               Tel: +55-11-2704922
               Fax: +55-11-2707001
               Email: asimpson@ludwig.org.br
               This sequence was derived from the FAPESP/LICR Human Cancer Genome
               Project. This entry can be seen in the following URL
               (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-CM4-FT0104-230
               600-215-c09&t3=2000-06-23&t4=1)
               Seq primer: puc 18 forward
               High quality sequence start: 25
               High quality sequence stop: 360.
               Location/Qualifiers
FEATURES       source
               1..361
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="FT0104"
               /dev_stage="Adult"
               /note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
               Site 2: SmaI; A mini-library was made by cloning products
               derived from ORESTES PCR (U.S. Letters Patent application
               No. 196,716 - Ludwig Institute for Cancer Research)
               profiles into the puc 18 vector. Reverse transcription of
               tissue mRNA and cDNA amplification were performed under
               low stringency conditions."
BASE COUNT    86 a 85 c 105 g 85 t
ORIGIN
Query Match   60.5%; Score 124.6; DB 12; Length 361;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAGAACGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTCTATTATCT 87
Db 258 AAGCGCAAGATCCAGTCAAGGTCAGTCTCCACTAAGCGTGGCTTGGCCCATATCT 199

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
Db 198 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGTGTGTTGAAGATACTGACTGCCAG 139

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 138 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTTTCGTTCCCGAGTGAGAG 80

RESULT 10
BE715602/c
LOCUS          BE715602
DEFINITION    CM4-HT0744-160600-201-a06 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION     BE715602
VERSION       BE715602.1 GI:10103867
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 364)
               Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
               Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
               Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
               Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
               , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
               Simpson, A.J.
TITLE          Shotgun sequencing of the human transcriptome with ORF expressed
               sequence tags
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE        20202663
COMMENT        Contact: Simpson A.J.G.
               Laboratory of Cancer Genetics
               Ludwig Institute for Cancer Research
               Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
               Brazil
               Tel: +55-11-2704922
               Fax: +55-11-2707001
               Email: asimpson@ludwig.org.br
               This sequence was derived from the FAPESP/LICR Human Cancer Genome
               Project. This entry can be seen in the following URL
               (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-CM4-HT0744-160
               600-201-a06&t3=2000-06-16&t4=1)
               Seq primer: puc 18 forward
               High quality sequence start: 20
               High quality sequence stop: 364.
               Location/Qualifiers
FEATURES       source
               1..364
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="HT0744"
               /dev_stage="Adult"
               /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
               Site 2: SmaI; A mini-library was made by cloning products
               derived from ORESTES PCR (U.S. Letters Patent application
               No. 196,716 - Ludwig Institute for Cancer Research)
               profiles into the puc 18 vector. Reverse transcription of
               tissue mRNA and cDNA amplification were performed under
               low stringency conditions."
BASE COUNT    85 a 92 c 100 g 87 t
ORIGIN
Query Match   60.5%; Score 124.6; DB 12; Length 364;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAGAACGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTCTATTATCT 87
Db 241 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCGCTCCCATATCT 182

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
Db 181 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGTGTGTTGAAGATACTGACTGCCAG 122

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 121 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTTTCGTTCCCGAGTGAGAG 63

RESULT 11
BE715605/c
LOCUS          BE715605
DEFINITION    CM4-HT0744-160600-201-a10 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION     BE715605
VERSION       BE715605.1 GI:10103870
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 370)

```

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4-HT0744-160

600-201-a10&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 370.

Location/Qualifiers

1. .370

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0744"

/dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 85 a 97 c 99 g 89 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 370;

Best Local Similarity 81.0%; Pred. No. 2.8e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

DB 242 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCT 183

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTCGCAG 147

DB 182 TGATCCGGTGGCGCATGTTGAATCCCGCTTAACCGCTGCTTGAAGATAGTACTGCTGCCAG 123

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTGTTTCGTTCCACAATAATAG 206

DB 122 GAATCAAGAAGTCTGTGAAGGCTCTTTCGGGATGGCCTGTTTCGTTCCCAAGTGAGAG 64

RESULT 12

BE715626

LOCUS

BE715626

CM4-HT0744-160600-201-h09 HT0744 Homo sapiens linear EST 12-SEP-2000

ACCESSION

BE715626

VERSION

BE715626.1 GI:10103891

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 370)

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

TITLE

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4-HT0744-160

600-201-h09&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 13

High quality sequence stop: 370.

Location/Qualifiers

1. .370

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0744"

/dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 85 a 110 c 87 g 88 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 370;

Best Local Similarity 81.0%; Pred. No. 2.8e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

DB 95 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCT 154

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTCGCAG 147

DB 155 TGATCCGGTGGCGCATGTTGAATCCCGCTTAACCGCTGCTTGAAGATAGTACTGCTGCCAG 214

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTGTTTCGTTCCACAATAATAG 206

DB 215 GAATCAAGAAGTCTGTGAAGGCTCTTTCGGGATGGCCTGTTTCGTTCCCAAGTGAGAG 273

RESULT 13

BE772160/c

LOCUS

BE772160

CM4-FT0104-230600-215-c08 FT0104 Homo sapiens linear EST 20-SEP-2000

ACCESSION

BE772160

VERSION

BE772160.1 GI:10225818

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 370)

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&tl2=IL5-HT0990-600-215-c08&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 370.

FEATURES

source 1..370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 86 a 95 c 99 g 90 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 370;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTCTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
DB 241 AAGCCAGAGAGCCAGTCAAGAGTCAGTCTCCACTAAGCCGTGGCTCTGCCCATTAATCT 182
QY 88 TGATTCGTTGGCCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATGTCAG 147
DB 181 TGATCCGTTGGCCATGTTGAATCCCTAACCGTCTGTTGAAAGATACACTGACTGCCAG 122
QY 148 GTATCAAAAGTCTGTCAGGTTCTCGGATGGCTGTTTCGTTCCCAATAATAG 206
DB 121 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 63

RESULT 14
BF838555
LOCUS IL5-HT0990-251100-263-g08 HT0990 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF838555
ACCESSION BF838555.1 GI:12190853
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 371)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE COMMENT

20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&tl2=IL5-HT0990-251100-263-g08&t3=2000-11-25&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 367.

FEATURES

source 1..371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0990"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 82 a 117 c 85 g 87 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 371;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTCTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
DB 75 AAGCCAGAGAGCCAGTCAAGAGTCAGTCTCCACTAAGCCGTGGCTCTGCCCATTAATCT 134
QY 88 TGATTCGTTGGCCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATGTCAG 147
DB 135 TGATCCGTTGGCCATGTTGAATCCCTAACCGTCTGTTGAAAGATACACTGACTGCCAG 194
QY 148 GTATCAAAAGTCTGTCAGGTTCTCGGATGGCTGTTTCGTTCCCAATAATAG 206
DB 195 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 253

RESULT 15
BE772166/c
LOCUS CM4-FT0104-230600-215-d05 FT0104 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE772166
ACCESSION BE772166.1 GI:10225824
VERSION
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 371)

REFERENCE
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM4-FT0104-230>)
 600-215-d05&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 371.
 Location/Qualifiers

FEATURES

source

1. .371
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"
 /dev_stage="Adult"

/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196.716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 84 a 96 c 101 g 90 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 371;
 Best Local Similarity 81.0%; Pred. No. 2.8e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87
 Db 242 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCCCCAATATCT 183
 QY 88 TGATTCGTCGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCAC 147
 Db 182 TGATCGCGTGGCGCATGTTGAATCCCGCTAACCGCTGCTTGAAGAGACTGACTGCCAC 123
 QY 148 GTATCAAAAGTCTGTGAGGTTCTCGGATGCTGTTTCGTTCCACATAATAG 206
 Db 122 GAATCAAGAAGTCTGTGAAGGCTCTTTCGGGATGCGCTGTTTCGTTCCCGGATGAG 64

RESULT 16

AI285400/c

LOCUS

DEFINITION

qt69f06.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:1960547 3'

similar to gb:S58717.rnal ELAFIN PRECURSOR (HUMAN); contains element

THR repetitive element ;, mRNA sequence.

ACCESSION

AI285400

VERSION

AI285400.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 375)

REFERENCE

AUTHORS

TITLE

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
 Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
 High quality sequence stop: 356.
 Location/Qualifiers

FEATURES

source

1. .375

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:1960547"

/clone_lib="NCI CGAP Eso2"

/tissue_type="squamous cell carcinoma"

/lab_host="DH10B"

/note="Organ: esophagus; Vector: pCMV-SPORT6; Site_1: SmaI

; Site_2: NotI; Cloned unidirectionally. Primer: Oligo

dr. Average insert size 1.1 kb. Life Technologies catalog

#: 11502-010"

BASE COUNT 91 a 83 c 125 g 75 t 1 others

ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 375;

Best Local Similarity 81.0%; Pred. No. 2.8e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTATTATCT 87

Db 351 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCAATATCT 292

QY 88 TGATTCGTCGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCAC 147

Db 291 TGATCCGTCGCGCATGTTGAATCCCGCTAACCGCTGCTTGAAGAGACTGACTGCCAC 232

QY 148 GTATCAAAAGTCTGTGAGGTTCTCGGATGCTGTTTCGTTCCACATAATAG 206

Db 231 GAATCAAGAAGTCTGTGAAGGCTCTTTCGGGATGCGCTGTTTCGTTCCCGGATGAG 173

RESULT 17

BE715603/c

LOCUS

DEFINITION

CM4-HT0744-160600-201-a07 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE715603

VERSION

BE715603.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 379)

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., D.H.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

MEDLINE

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM4-HT0744-160>)

600-201-a07&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 25

High quality sequence stop: 379.

Location/Qualifiers

1. .379

BASE COUNT 86 a 95 c 112 g 86 t
 ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 379;
 Best Local Similarity 81.0%; Pred. No. 2.9e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTAAAGTCTGCTGCTTACTAAGCCAGGTTCTGTCTATTATCT 87
 DB 262 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 203

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAG 147
 DB 202 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGACGCCCCAG 143

QY 148 GTATCAAAAGTGTGTTGAAGTTCCTCGGTATGGCTTGTTCGTTCCCAATAATAG 206
 DB 142 GAATCAAGAGTGTGTTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCAAGTGAGAG 84

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

/dev_stage="Adult"
 /note="Organ: prostate tumor; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 89 a 99 c 100 g 92 t
 ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 380;
 Best Local Similarity 81.0%; Pred. No. 2.9e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTAAAGTCTGCTGCTTACTAAGCCAGGTTCTGTCTATTATCT 87
 DB 249 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 190

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAG 147
 DB 189 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATAGTACTGACGCCCCAG 130

QY 148 GTATCAAAAGTGTGTTGAAGTTCCTCGGTATGGCTTGTTCGTTCCCAATAATAG 206
 DB 129 GAATCAAGAGTGTGTTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCAAGTGAGAG 71

RESULT 19
 BE772171/c
 LOCUS BE772171 380 bp mRNA linear EST 20-SEP-2000
 DEFINITION CM4-FT0104-230600-215-g07 FT0104 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE772171
 VERSION BE772171.1 GI:10225829
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 380)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
 600-215-b05&t3=2000-06-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 365.

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

FEATURES
 source
 Location/Qualifiers
 1..380
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FT0104"

derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 89 a 97 c 106 g 88 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 380;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCGATTAAAGTCTCTGTCTACTAAGCCAGGTCTTGTCTCTATTATCT 87
DB 253 AAGCGCAAGAGGCGAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 194

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTCGCCAG 147
DB 193 TGATCGGTGGCGCATGTTGAATCTCCCTTAACCGTGTGTTGAAGATCTGACTGCCAG 134

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCCTCCACAATAATAG 206
DB 133 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTTGTTCCTCCCACTGAGAG 75

RESULT 20
BE182879
LOCUS BE182879 386 bp mRNA linear EST 22-JUN-2000
DEFINITION CM4-HT0652-150400-143-g08 HT0652 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE182879
VERSION BE182879.1 GI:8662055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldmann, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=CM4-HT0652-150400-143-g08&t3=2000-04-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 61.
Location/Qualifiers
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under

FEATURES
source
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under

BASE COUNT 89 a 104 c 105 g 88 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 10; Length 386;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCGATTAAAGTCTCTGTCTACTAAGCCAGGTCTTGTCTCTATTATCT 87
DB 144 AAGCGCAAGAGGCGAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 203

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTCGCCAG 147
DB 204 TGATCGGTGGCGCATGTTGAATCTCCCTTAACCGTGTGTTGAAGATCTGACTGCCAG 263

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCCTCCACAATAATAG 206
DB 264 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCTTGTTCCTCCCACTGAGAG 322

RESULT 21
BE772169/c
LOCUS BE772169 386 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-F0104-230600-215-f02 F0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772169
VERSION BE772169.1 GI:10225827
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldmann, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=CM4-F0104-230600-215-f02&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 45
High quality sequence stop: 386.
Location/Qualifiers
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="F0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under

FEATURES
source
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="F0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under

BASE COUNT 90 a 99 c 107 g 90 t
ORIGIN

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 390)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-c07&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 390.

FEATURES
source
1..390
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No.196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 90 a 97 c 111 g 92 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 12; Length 390;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Qy 28 ATGCTCAAGAACCCAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGCTTATATCT 87
Db 261 AAGCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCGTGGCTCTGCCCATTAATCT 202
Qy 88 TGATTCGTTGGCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATGTCCAG 147
Db 201 TGATCCGTTGGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAGAGATACTGACTGCCAG 142
Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGCGTATGGCTGTTTCGTTCCACAATAATAG 206
Db 141 GAATCAAGAGTGTGTGAAGGCTCTTGC GGATGGCTGTTTCGTTCCCGCAGTGAGAG 83

RESULT 27
BE715615
LOCUS BE715615 395 bp mRNA linear EST 12-SEP-2000
DEFINITION CM4-HT0744-160600-201-d10 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715615
VERSION BE715615.1 GI:10103880
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 395)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-d10&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 395.
Location/Qualifiers
1..395
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No.196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 91 a 115 c 96 g 93 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 12; Length 395;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Qy 28 ATGCTCAAGAACCCAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGCTTATATCT 87
Db 120 AAGCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCGTGGCTCTGCCCATTAATCT 179
Qy 88 TGATTCGTTGGCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATGTCCAG 147
Db 180 TGATCCGTTGGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAGAGATACTGACTGCCAG 239
Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGCGTATGGCTGTTTCGTTCCACAATAATAG 206
Db 240 GAATCAAGAGTGTGTGAAGGCTCTTGC GGATGGCTGTTTCGTTCCCGCAGTGAGAG 298

RESULT 28

BE772137

LOCUS

BE772137

CM4-F70104-230600-215-a12 FT0104 Homo sapiens cDNA, mRNA sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 397)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230>

600-215-al2&t3=2000-06-23&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 4

High quality sequence stop: 397.

FEATURES

source

1. .397

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="FT0104"

/dev_stage="Adult"

/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 397;

Best Local Similarity 81.0%; Pred. No. 2.9e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTCTATTATCT 87

Db 122 AAGCGCAAGAGCCAGTCAAAGGCCAGTCCACTCAAGCTGCTCTGCCCATTTATCT 181

QY 88 TGATTCGTCGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147

Db 182 TGATCCGGTCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAGATCTGACTGCCAG 241

QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCGCGGTATGCTTTCGTTCCCAATATAG 206

Db 242 GAATCAAGAAGTGTGTGAAGGCTTTGCGGGATGCGCTGTTCTGCCAGTGAGAG 300

RESULT 29

AA586974/c

LOCUS

DEFINITION

nm9608.s1 NCI CGAP Lari Homo sapiens cDNA clone IMAGE:1089135 3'

similar to gb:S58717.rnai ELAFIN PRECURSOR (HUMAN);contains element

THR repetitive element ;, mRNA sequence.

AA586974

VERSION

AA586974.1 GI:2397788

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 399)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

cDNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 523 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham.

Location/Qualifiers

1. .399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCI CGAP Lari"

/tissue_type="larynx"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: larynx; Vector: Bluescript SK-; Site 1:

EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:

Oligo df. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGG

3' 3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTT 3'

Average insert size: 0.9 kb."

97 a 88 c 129 g 85 t

Query Match 60.5%; Score 124.6; DB 9; Length 399;

Best Local Similarity 81.0%; Pred. No. 2.9e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTATTATCT 87

Db 351 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAAGCTGCTCTGCCCATTTATCT 292

QY 88 TGATTCGTCGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147

Db 291 TGATCCGGTCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAGATCTGACTGCCAG 232

QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCGCGGTATGCTTTCGTTCCCAATAATAG 206

Db 231 GAATCAAGAAGTGTGTGAAGGCTTTGCGGGATGCGCTGTTCTGCCAGTGAGAG 173

RESULT 30

BE715592

LOCUS

DEFINITION

CM4-HT0744-160600-201-b10 HT0744 Homo sapiens cDNA, mRNA sequence.

BE715592

ACCESSION

BE715592.1 GI:10103857

VERSION

EST.

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 400)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

TITLE

JOURNAL

MEDLINE

COMMENT

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=CM4-HT0744-160
 600-201-b10at3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 7
 High quality sequence stop: 400.
 Location/Qualifiers

FEATURES

source
 1. .400
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 90 a 115 c 100 g 95 t

Query Match 60.5%; Score 124.6; DB 12; Length 400;
 Best Local Similarity 81.0%; Pred. No. 2.9e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCGTAAAGTCTGTCTACTAAGCCAGGTTCTTGCTTATTAATCT 87
 |||
 Db 121 AAGCCAAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCTGGCTCGGCCATTAATCT 180
 |||
 Qy 88 TGATTCTGTTGGCTATGTTAAACCCACCTAACCTGTTTTCGAAGGACACTGATTGTCCAG 147
 |||
 Db 181 TGATCCGGTGGCCATGTTGAATCCCTCAACCGTGTCTGAAGATAGTACTGACTGCCAG 240
 |||
 Qy 148 GTATCAAAAGTGCTGTGAAGTTCCTCGGATGCTGTTTTCGTTCCACAATAATAG 206
 |||
 Db 241 GAATCAAGAAGTGCTGTGAAGCTCTTCGGGATGGCTGTTTCGTTCCCGCAGTGAGAG 299
 |||

RESULT 31
 BE715609/c
 LOCUS BE715609 403 bp mRNA linear EST 12-SEP-2000
 DEFINITION CM4-HT0744-160600-201-c04 HT0744 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE715609
 VERSION BE715609.1 GI:10103874
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 403)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsuoka,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=CM4-HT0744-160
 600-201-c04at3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 12
 High quality sequence stop: 403.
 Location/Qualifiers

FEATURES

source
 1. .403
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"

/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 96 a 98 c 113 g 96 t

Query Match 60.5%; Score 124.6; DB 12; Length 403;
 Best Local Similarity 81.0%; Pred. No. 2.9e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCGTAAAGTCTGTCTACTAAGCCAGGTTCTTGCTTATTAATCT 87
 |||
 Db 274 AAGCCAAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCTGGCTCGGCCATTAATCT 215
 |||
 Qy 88 TGATTCTGTTGGCTATGTTAAACCCACCTAACCTGTTTTCGAAGGACACTGATTGTCCAG 147
 |||
 Db 214 TGATCCGGTGGCCATGTTGAATCCCTCAACCGTGTCTGAAGATAGTACTGACTGCCAG 155
 |||
 Qy 148 GTATCAAAAGTGCTGTGAAGTTCCTCGGATGCTGTTTTCGTTCCACAATAATAG 206
 |||
 Db 154 GAATCAAGAAGTGCTGTGAAGCTCTTCGGGATGGCTGTTTCGTTCCCGCAGTGAGAG 96
 |||

RESULT 32
 AI874186/c
 LOCUS AI874186 405 bp mRNA linear EST 07-MAR-2000
 DEFINITION wms0a02.xl NCI CGAP Ut4 Homo sapiens cDNA clone IMAGE:2439338 3',
 similar to gb:S58717_rnal ELAPIN PRECURSOR (HUMAN);, mRNA sequence.
 ACCESSION AI874186
 VERSION AI874186.1 GI:5548235
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 405)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 644 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES

source
 1. .405
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

```

/clone="IMAGE:2439338"
/clone_lib="NCI CGAP Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"
BASE COUNT 96 a 84 c 125 g 100 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 9; Length 405;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACGCTTAAAGCTCTGCTCTACTAAGCCAGTTCCTCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTCGCTCTGCCCATTAATCT 318
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGGACACTGATTCCTCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 TGATCCGGTGGCCAGTGTGAATCCCTTAACCGTCTTGAAGATACTGACTGCCAG 258
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCCTCACAAATATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 GAATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCGCTGTTTCGTTCCCGAGTGAAG 199
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 33
AW137392/c
LOCUS
DEFINITION
IMAGE:2715238 3', mRNA sequence.
AW137392
VERSION
EST.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 406)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LJNL at:
www-bio.ljnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers
1. 406
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2715238"
/clone_lib="NCI CGAP Sub3"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NCI-CGAP Sub3 library is a subtracted library derived from
the NCI-CGAP Sub1 library, which is a subtracted library
derived from BI. BI constitutes a mixture of 21
normalized or subtracted NCI CGAP libraries: NCI-CGAP_Co4
, NCI-CGAP_Pr22, NCI-CGAP_Pr28, NCI-CGAP_Co10,
NCI-CGAP_Co16, NCI-CGAP_Kid5, NCI-CGAP_Kid12,
NCI-CGAP_Kid3, NCI-CGAP_Kid11, NCI-CGAP_Lym2,

```

FEATURES

source

```

NCI-CGAP_Br2, NCI-CGAP_Co8, NCI-CGAP_CLL1, NCI-CGAP_Lei2,
NCI-CGAP_Brn23, NCI-CGAP_Lu5, NCI-CGAP_Lu24,
NCI-CGAP_Lu19, NCI-CGAP_GC4, NCI-CGAP_GC6,
NCI-CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI-CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clones 132376-132391, 1456008-1456775
, 1500552-1502855); NCI-CGAP_Kid5 pool 1 LLAM 3338-3342
, 3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831,
1471368-1472903, 1492104-1493255); NCI-CGAP_Lu5 pool 1
LLAM 3575-3582, 3851-3854 (IMAGE Clones 1414920-1417991,
1520904-1522439); NCI-CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE Clones 1257096-1258631,
146064-1470983, 1475592-1476743); NCI-CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones
985608-986759, 1101192-1101959, 1217928-1220615);
NCI-CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
Clones 1057416-1061255, 114584-1145311). Subtraction
was performed as previously described (Bonaldi, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.
TAG_LIB=NCI-CGAP_Co10
TAG_TISSUE=Colon
TAG_SEQ=AAACG"
BASE COUNT 95 a 86 c 126 g 99 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 10; Length 406;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACGCTTAAAGTCTCTGCTCTACTAAGCCAGTTCCTCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 368 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTCGCTCTGCCCATTAATCT 309
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGGACACTGATTCCTCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 308 TGATCCGGTGGCCAGTGTGAATCCCTTAACCGTCTTGAAGATACTGACTGCCAG 249
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCCTCACAAATATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 GAATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCGCTGTTTCGTTCCCGAGTGAAG 190
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 34
BE715604
LOCUS
DEFINITION
CM4-HT0744-160600-201-a08 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE715604
VERSION
BE715604.1 GI:10103869
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 406)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Grundman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

```

Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=8t2=CM4-HT0744-160
 600-201-a08&t3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 406.
 Location/Qualifiers

FEATURES

1..406
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORSTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 94 a 118 c 97 g 97 c

BASE COUNT

ORIGIN
 Query Match 60.5%; Score 124.6; DB 12; Length 406;
 Best Local Similarity 81.0%; Pred. No. 2.9e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
 |||||
 Db 131 AAGCGCAAGAGCGCTCAAGAGTCCAGTCTCCACTAAGCGTGGCTCTGCCCATTTATCT 190
 |||||
 QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
 |||||
 Db 191 TGATCGGTGCCCATGTTGAATCCCTTAACCGTCTTGAAGATATCTGACTGCCAG 250
 |||||
 QY 148 GTATCAAAAGTGTGTGAAGTTCCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
 |||||
 Db 251 GAATCAAGAAGTGTGTGAAGCTCTTGGGGATGGCTTGTTCGTTCCCGCAGTGAGAG 309
 |||||

RESULT 35

AA586943/c
 LOCUS
 DEFINITION
 AA586943 419 bp mRNA linear EST 26-SEP-1997
 nm69a01.s1 NCI CGAP Lax1 Homo sapiens cDNA clone IMAGE:1089096 3',
 similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
 HGR repetitive element ;, mRNA sequence.
 AA586943
 VERSION
 AA586943.1 GI:2397757
 KEYWORDS
 EST.
 SOURCE
 human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 419)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 615 Std Error: 0.00

Seq primer: -40mi3 fwd. ET from Amersham
 High quality sequence stop: 335.
 Location/Qualifiers

FEATURES

1..419
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1089096"
 /clone_lib="NCI CGAP Lax1"
 /tissue_type="larynx"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: larynx; Vector: Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Larynx. 5' adaptor sequence: 5' GAATTCGACGAG
 3' 3' adaptor sequence: 5' CTCAGATTTTTTTTTTTTTTTT 3'
 Average insert size: 0.9 kb."
 103 a 92 c 135 g 89 c

BASE COUNT

ORIGIN
 Query Match 60.5%; Score 124.6; DB 9; Length 419;
 Best Local Similarity 81.0%; Pred. No. 3e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
 |||||
 Db 351 AAGCGCAAGAGCGCTCAAGAGTCCAGTCTCCACTAAGCGTGGCTCTGCCCATTTATCT 292
 |||||
 QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
 |||||
 Db 291 TGATCGGTGCCCATGTTGAATCCCTTAACCGTCTTGAAGATATCTGACTGCCAG 232
 |||||
 QY 148 GTATCAAAAGTGTGTGAAGTTCCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
 |||||
 Db 231 GAATCAAGAAGTGTGTGAAGCTCTTGGGGATGGCTTGTTCGTTCCCGCAGTGAGAG 173
 |||||

RESULT 36

AA082097/c
 LOCUS
 DEFINITION
 AA082097 422 bp mRNA linear EST 14-OCT-1999
 xb60d08.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:2580687 3',
 similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
 MER28 MER28 repetitive element ;, mRNA sequence.
 AA082097
 VERSION
 AA082097.1 GI:6037249
 KEYWORDS
 EST.
 SOURCE
 human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 422)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

AUTHORS

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
 Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

JOURNAL

COMMENT
 Possible reversed clone: polyt not found
 Seq primer: -40UP from Gibco
 High quality sequence stop: 407.
 Location/Qualifiers
 1..422
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2580687"
 /clone_lib="NCI CGAP_Eso2"

/tissue_type="squamous cell carcinoma"
 /lab_host="DH108"
 /note="Organ: esophagus; Vector: pCMV-SPORT6; Site 1: Sali
 ; Site 2: NotI; Cloned unidirectionally. Primer: Oligo
 dr. Average insert size 1.1 kb. Life Technologies catalog
 #: 11502-010"

BASE COUNT 98 a 88 c 129 g 107 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 10; Length 422;

Best Local Similarity 81.0%; Pred. No. 3e-29; Mismatches 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 253 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 195

RESULT 37

BE772154/c

LOCUS

DEFINITION CM4.FT0104-230600-215-b04 FT0104 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE772154

VERSION BE772154.1

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 428)

BE772154

CM4.FT0104-230600-215-b04 FT0104 Homo sapiens cDNA, mRNA sequence.

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 253 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 195

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 253 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 195

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 253 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 195

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 253 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 195

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 253 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 195

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 373 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 314

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 313 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTGAAGATACACTGACTGCCAG 254

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 101 a 99 c 131 g 97 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 428;

Best Local Similarity 81.0%; Pred. No. 3e-29; Mismatches 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

RESULT 38

AI283910/c

LOCUS

DEFINITION

AI283910

VERSION

AI283910

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 429)

AI283910

AI283910

VERSION

AI283910

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 429)

AI283910

AI283910

VERSION

AI283910

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 429)

AI283910

AI283910

VERSION

AI283910

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Query Match 60.5%; Score 124.6; DB 12; Length 428;

Best Local Similarity 81.0%; Pred. No. 3e-29; Mismatches 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAAATAAG 206

Db 215 GAATCAAGAAGTCTGTGAAGGCTTTGCGGATGCGCTTTCGTTCCCGAGTGAGAG 157

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87

Db 335 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTCCGCCATTATCT 276

QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATTGTCAG 147

Db 275 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGCTGTTCGTTGAAGATACACTGACTGCCAG 216

QY 148 GTATCAAAAAGTCTGTGA

[illegible][illegible]

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 437)

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bail, G. S., Simpson, D. H., Brunstein, A., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.

```

FEATURES
  source
    high quality sequence stop: 437.
    Location/Qualifiers
      1. 437
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_lib="HT1016"
        /dev stage="Adult"
        /note="Organ: head neck; Vector: puc18; Site_1: SmaI;
        Site_2: SmaI; A mini-library was made by cloning products
        derived from ORESTS PCR (U.S. Letters Patent application
        No. 196,716 - Ludwig Institute for Cancer Research)
        profiles into the pUC 18 vector. Reverse transcription of
        tissue mRNA and cDNA amplification were performed under
        low stringency conditions."
      99 a 120 c 118 g 100 t

```

[illegible][illegible]

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 438)

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

```

FEATURES
Source
Location/Qualifiers
i. .438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 102 c 132 g 100 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 12; Length 438;
Best Local Similarity 81.0%; Pred. No. 3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```


Db 269 TGATCGGTCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGACTAGTACCTGCCAG 210
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCGTTCCACAAATAATAG 206
Db 209 GAATCAAGAAGTCTGTGAAGGCTTTCGCGGATGCGCTTTCGTTCCCAAGTGAAG 151

RESULT 41
BF836743
LOCUS
DEFINITION
CM2-HT0969-181100-509-b06 HT0969 Homo sapiens cDNA, mRNA sequence.
BF836743
VERSION
BF836743.1 GI:12188790
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 441)
Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM2&t2=CM2-HT0969-
181100-509-b06&t3=2000-11-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 441.

FEATURES
source
1. 441
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0969"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 131 c 109 g 107 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 441;
Best Local Similarity 81.0%; Pred. No. 3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTAAAGGCTCTGTCTACTAAGCCAGGTTCTTGTCTATTATCT 87
Db 146 AAGCGCAAGAGCAGTCAAGAGGCTCAAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 205
QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 206 TGATCCGGTGGCAATGTTGAATCCCGCTAACCGCTGTTGAAAGATATCTAGTCCGCCAG 265
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCGCGTATGCTTGTTCGTTCCACAAATAATAG 206

Db 266 GAATCAAGAAGTCTGTGAAGGCTTTCGCGGATGCGCTTTCGTTCCCAAGTGAAG 324

RESULT 42
AI582329/c
LOCUS
DEFINITION
AI582329.1
ACCESSION
AI582329
VERSION
AI582329.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 448)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@emil.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40UP from Gibco
POLYA-No.

FEATURES
source
1. 448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2213824"
/clone_lib="NCI_CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="Adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from a
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 105 a 96 c 138 g 109 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 448;
Best Local Similarity 81.0%; Pred. No. 3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTAAAGGCTCTGTCTACTAAGCCAGGTTCTTGTCTATTATCT 87
Db 366 AAGCGCAAGAGCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 307
QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAG 147
Db 306 TGATCCGGTGGCCATGTTGAATCCCGCTAACCGCTCTTGAAGAGATCTGACTGCCAG 247
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCGCGTATGCTTGTTCGTTCCACAAATAATAG 206
Db 246 GAATCAAGAAGTGTGTGAAGGCTCTTTCGCGGATGCGCTTTCGTTCCCAAGTGAAG 188

RESULT 43
 AI459240/c
 LOCUS
 DEFINITION tkilc09.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2150704 3' similar to gb:558717_rnal ELAFIN PRECURSOR (HUMAN); contains element THR repetitive element ;, mRNA sequence.
 ACCESSION AI459240
 VERSION AI459240.1 GI:43111819
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 453)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 cDNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 554 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES

Location/Qualifiers
 1..453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2150704"
 /clone_lib="NCI CGAP Lu24"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Ronaldo."

BASE COUNT 107 a 96 c 137 g 113 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 9; Length 453;
 Best Local Similarity 81.0%; Pred. No. 3e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTTATATCT 87
 Db 368 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCCTGGCTCTGCCCATTTATCT 309
 QY 88 TGATTCTGGTGGCTATGTTAAACCCACTAACCCCTGTTTGAAGGACACTGATTGTCCAG 147
 Db 308 TGATCCGGTGGCCATGTTGTAATCCCTTAACCCGCTGCTTGAAGATAGTACTGCTGCCAG 249
 QY 148 GTATCAAAAAGTGTGGAAGGTTCTTGGGATGCGGTATGGCTTGTTCGTTCCCAATAATAG 206
 Db 248 GAATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCGCAGTGAGAG 190

RESULT 44
 BF824934
 LOCUS
 DEFINITION ILO-HN0039-101100-500-b06 HN0039 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF824934
 VERSION BF824934.1 GI:12167001
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 465)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldmann, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., Geoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL06t2=ILO-HN0039-101100-500-b06&t3=2000-11-10&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 465.

FEATURES

Location/Qualifiers
 1..465
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

98 a 137 c 117 g 113 t

Query Match 60.5%; Score 124.6; DB 12; Length 465;
 Best Local Similarity 81.0%; Pred. No. 3.1e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTTATATCT 87
 Db 181 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCCTGGCTCTGCCCATTTATCT 240
 QY 88 TGATTCTGGTGGCTATGTTAAACCCACTAACCCGTTGTTTGAAGGACACTGATTGTCCAG 147
 Db 241 TGATCCGGTGGCCCATGTTGAATCCCTCAACCCGCTGCTTGAAGAGATACTGACTGCCAG 300
 QY 148 GTATCAAAAAGTGTGGAAGGTTCTTGGGATGCGGTATGGCTGTTTCGTTCCCAATAATAG 206
 Db 301 GAATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCGCAGTGAGAG 359

RESULT 45
 BF825207/c
 LOCUS
 DEFINITION ILO-HN0039-131100-501-d07 HN0039 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF825207
 VERSION BF825207.1 GI:12167451
 KEYWORDS EST.
 SOURCE human.

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 467)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?Cl=IL0&t2=IL0-HN0039-
131100-501-d07&t3=2000-11-13&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 467.
FEATURES
source
Location/Qualifiers
1..467
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 115 c 137 g 99 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 12; Length 467;
Best Local Similarity 81.0%; Pred. No. 3.1e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGCTATTATCT 87
Db 285 AAGCGCAAGAGCCAGTCAAAAGTCCAGTCTCCACTAAGCGTGGCTCCTGCCCATTTATCT 226
QY 88 TGATTCGGTCCGCTATGTTAAACCCACCTAACCGTTGTTGAGGACACTGATGTCAG 147
Db 225 TGATTCGGTCCGCTATGTTAAACCCACCTAACCGTTCGTTGAAAGATATCTGATGCCAG 166
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCGCGGTATGCTGTTTCCATCAATATAG 206
Db 165 GAATCAAGAAGTGTGTGAAGGCTCTTCGCGGATGCGCTGTTCTCCAGTGAGAG 107
RESULT 46
BF837734/c 471 bp mRNA linear EST 13-JAN-2001
LOCUS BF837734
DEFINITION QV3-HT1016-221100-480-g03 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF837734
VERSION BF837734.1 GI:12189956
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 471)

```

```

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?Cl=QV3&t2=QV3-HT1016-
221100-480-g03&t3=2000-11-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 468.
FEATURES
source
Location/Qualifiers
1..471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 114 c 138 g 102 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 12; Length 471;
Best Local Similarity 81.0%; Pred. No. 3.1e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGCTATTATCT 87
Db 301 AAGCGCAAGAGCCAGTCAAAAGTCCAGTCTCCACTAAGCGTGGCTCCTGCCCATTTATCT 242
QY 88 TGATTCGGTCCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCAG 147
Db 241 TGATCCGGTGGCCATGTTGAATCCCGCTAACCGCTGCTTGAAGATACCTGACTGCCAG 182
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCGCGGTATGCTGTTTCCATCAATATAG 206
Db 181 GAATCAAGAAGTGTGTGAAGGCTCTTCGCGGATGCGCTGTTCTCCAGTGAGAG 123
RESULT 47
BF824932 477 bp mRNA linear EST 13-JAN-2001
LOCUS BF824932
DEFINITION IL0-HN0039-101100-500-a12 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824932
VERSION BF824932.1 GI:12166998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 477)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

```

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-HN0039-101100-500-412&t3=2000-11-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 476.

FEATURES

source
 1. 477
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 103 a 139 c 121 g 114 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 12; Length 477;
 Best Local Similarity 81.0%; Pred. No. 3.1e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 Qy 28 ATGCTCAAGAACCAAGTCTGTCTACTAAGCCAGGTTCTGTCTCTATATCT 87
 Db 182 AAGCGAAGACCAAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATATCT 241
 Qy 88 TGATTCGTGGCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATGTCAG 147
 Db 242 TGATCCGTGGCCCATGTTGAATCCCTTACCCGCTGCTTGAAGATACCTGACCTGCCAG 301
 Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGCGTATGGCTTGTTCGTTCCCAATAATAG 206
 Db 302 GAATCAAGAAAGTGTGTAAGGCTCTTGCAGGATGGCTGTTTCGTTCCCAAGTGAGAG 360

RESULT 48
 AA582866/c
 LOCUS 480 bp mRNA linear EST 26-SEP-1997
 DEFINITION hn72a09.81 NCI_CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089400 3' similar to gb:858717.1nal ELAFIN PRECURSOR (HUMAN); contains element HGR repetitive element ; mRNA sequence.
 ACCESSION AA582866
 VERSION AA582866.1 GI:2360226
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 480)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 647 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 430.

FEATURES

source
 1. 480
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1089400"
 /clone_lib="NCI_CGAP_Lar1"
 /tissue_type="larynx"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: larynx; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGCGCAGCAG 3', 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3', Average insert size: 0.9 kb."

BASE COUNT 113 a 107 c 144 g 116 t
 ORIGIN
 Query Match 60.5%; Score 124.6; DB 9; Length 480;
 Best Local Similarity 81.0%; Pred. No. 3.1e-29;
 Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 Qy 28 ATGCTCAAGAACCAAGTCTGTCTACTAAGCCAGGTTCTGTCTCTATATCT 87
 Db 365 AAGCGAAGACCAAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATATCT 306
 Qy 88 TGATTCGTGGCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATGTCAG 147
 Db 305 TGATCCGTGGCCCATGTTGAATCCCTTACCCGCTGCTTGAAGATACCTGACCTGCCAG 246
 Qy 148 GTATCAAAAGTGTGTAAGGTTCTCGCGTATGGCTTGTTCGTTCCCAATAATAG 206
 Db 245 GAATCAAGAAAGTGTGTAAGGCTCTTGCAGGATGGCTGTTTCGTTCCCAAGTGAGAG 187

RESULT 49
 BF002099/c
 LOCUS 490 bp mRNA linear EST 06-OCT-2000
 DEFINITION 7G99b04.x1 NCI_CGAP Col6 Homo sapiens cDNA clone IMAGE:3314575 3' similar to SW:ELAF_HUMAN P19957 ELAFIN PRECURSOR ; contains element MER28 repetitive element ; mRNA sequence.
 ACCESSION BF002099
 VERSION BF002099.1 GI:10702374
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 490)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seq primer: -40UP from Gibco

```

High quality sequence stop: 421.
Location/Qualifiers
1. .490
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3314575"
/clone_lib="NCI_CGAP_Col6"
/tissue_type="colon tumor, RER+"
/lab_host="DH10B"
/note="Organ: colon; Vector: p7T73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not J; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351)
Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 119 a 113 c 146 g 112 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 490;
Best Local Similarity 81.0%; Pred. No. 3.1e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCGATTAAAGTCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87
Db 357 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 298

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
Db 297 TGATCCGGTGCAGCATGTTGAATCCCTAACCGTGTGTTGAAGATACCTGATGCCAG 238

QY 148 GTATCAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCGTTCCACATATAG 206
Db 237 GAATCAAGAAGTCTGTGAAGGCTCTTGGGATGCGCTTTCGTTCCCGCAGTGAGAG 179

RESULT 50
BF837735 498 bp mRNA linear EST 13-JAN-2001
LOCUS
DEFINITION QV3-HT1016-221100-480-g08 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF837735
VERSION BF837735.1 GI:12189957
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 498)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baig,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV3&t2=QV3-HT1016-
221100-480-g08&t3=2000-11-22&t4=1)
Seq primer: puc 18 forward

```

```

High quality sequence stop: 498.
Location/Qualifiers
1. .498
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 136,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 146 c 128 g 120 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 498;
Best Local Similarity 81.0%; Pred. No. 3.1e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCGATTAAAGTCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87
Db 201 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 260

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
Db 261 TGATCCGGTGCAGCATGTTGAATCCCTAACCGTGTGTTGAAGATACCTGATGCCAG 320

QY 148 GTATCAAAAGTCTGTGAAGGTTCTCGCGTATGCGTGTGTTGTTGTTCCACAATAAG 206
Db 321 GAATCAAGAAGTCTGTGAAGGCTCTTGGGATGCGCTGTTTCGTTCCCGCAGTGAGAG 379

RESULT 51
BQ941085 522 bp mRNA linear EST 21-AUG-2002
LOCUS
DEFINITION AGENCOURT 8794612 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374832
5', mRNA sequence.
ACCESSION BQ941085
VERSION BQ941085.1 GI:22356563
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 522)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs@mail.nih.gov
Tissue procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2554 row: h column: 01
High quality sequence stop: 521.
Location/Qualifiers
1. .522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6374832"
/clone_lib="NIH MGC 18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the

```

laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

```

BASE COUNT      124 a 155 c 121 g 122 t
ORIGIN
Query Match      60.5%; Score 124.6; DB 14; Length 522;
Best Local Similarity 81.0%; Pred. No. 3.2e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTTATATCT 87
Db 154 AAGCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCCTGCTGCTGCCCATTTATCT 213

Qy 88 TGATTCTGTGGCTATGTTAAACCCACCTAACCGTTCTTTGAAGGACACTGATTGTCCAG 147
Db 214 TGATCCGGTGGCCATGTTGAATCCCTAACCGTCTGTTGAAGATAGTACTGCTGCCAG 273

Qy 148 GTATCAAAAAGTGCTGTGAAGGTTCTTCGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 274 GAATCAAGAAGTCTGTGAAGGCTTTGCGGGATGGCTGTTTCGTTCCCGCAGTGAGAG 332

RESULT 52
BG490472      565 bp      mRNA      linear      EST 27-MAR-2001
LOCUS      602519610F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4637906 5',
DEFINITION      mRNA sequence.
ACCESSION      BG490472
VERSION      BG490472.1 GI:13451982
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 565)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI399 row: d column: 03
High quality sequence stop: 564.
FEATURES
source
1. .565
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4637906"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

BASE COUNT      141 a 163 c 130 g 131 t
ORIGIN
Query Match      60.5%; Score 124.6; DB 12; Length 565;
Best Local Similarity 81.0%; Pred. No. 3.3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 30 GCTCAAGAACCAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTTATATCTTG 89
Db 349 GCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCCTGCTGCTGCCCATTTATCTTG 290

```

```

Qy 28 ATGCTCAAGAACCAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTTATATCT 87
Db 186 AAGCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCCTGCTGCTGCCCATTTATCT 245

Qy 88 TGATTCTGTGGCTATGTTAAACCCACCTAACCGTTCTTTGAAGGACACTGATTGTCCAG 147
Db 246 TGATCCGGTGGCCATGTTGAATCCCTAACCGTCTGTTGAAGATAGTACTGACTGCCAG 305

Qy 148 GTATCAAAAAGTGCTGTGAAGGTTCTTCGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 306 GAATCAAGAAGTCTGTGAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCGCAGTGAGAG 364

RESULT 53
AI392753/c      418 bp      mRNA      linear      EST 30-MAR-1999
LOCUS      t923g03.x1 NCI_CGAP_C111 Homo sapiens cDNA clone IMAGE:2109652 3',
DEFINITION      similar to gb:587117_rnal ELAFIN PRECURSOR (HUMAN); contains element
HGR repetitive element ;, mRNA sequence.
ACCESSION      AI392753
VERSION      AI392753.1 GI:4222300
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 418)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 609 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 411.
FEATURES
source
1. .418
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:2109652"
/clone_lib="NCI_CGAP_C111"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGGCGGCGGATGCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT      103 a 92 c 134 g 88 t
ORIGIN
Query Match      60.3%; Score 124.2; DB 9; Length 418;
Best Local Similarity 81.4%; Pred. No. 4e-29;
Matches 144; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 30 GCTCAAGAACCAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTTATATCTTG 89
Db 349 GCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCCTGCTGCTGCCCATTTATCTTG 290

```


RESULT 56
BE711100
LOCUS RC3-HT0649-270700-012-f03 HT0649 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE711100
ACCESSION BE711100
VERSION BE711100.1 GI:1009365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 263)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC3-HT0649-270700-012-f03&t3=2000-07-27&t4=1)
700-012-f03&t3=2000-07-27&t4=1)
Seg primer: puc 18 forward
High quality sequence start: 50
High quality sequence stop: 263.
Location/Qualifiers
1. .263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
62 a 71 c 74 g 56 t

FEATURES

source
1. .263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
62 a 71 c 74 g 56 t

BASE COUNT

ORIGIN
Query Match 59.7%; Score 123; DB 12; Length 263;
Best Local Similarity 76.9%; Pred. No. 8.2e-29;
Matches 150; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 12 GGTACCTACTGTCATATGCTCAAGAACCAAGTCTGCTTAAAGTCTGCTTAAAGCCAGGT 71
Db 39 GGTGAGTAAGAGTCAAGCGCAAGAGCCAGTCAAGGTCCAGTCTCACTGAGCCCTGGC 98
Qy 72 TCTTGCTCTATTACTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAG 131
Db 99 TCCTGCCCATATCTTGATCCGTTGGCCCAATGTTGAATCCCTTAACCGCTGCTTGAAG 158
Qy 132 GACACTGATTCAGGATCAAAAGTCTGTTGAAGGTTCTCGCGGTATGGCTTGTTC 191
Db 159 GATACTGACTGCCAGGAATCAAGAGTCTGTTGAAGGCTCTTCGCGGATGGCTGTTTC 218
Qy 192 GTTCCACATAATAG 206
Db 219 GTACCCCAATGAGAG 233

RESULT 57

BE181373/c
LOCUS CM0-HT0633-240300-304-g12 HT0633 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE181373
ACCESSION BE181373
VERSION BE181373.1 GI:8660549
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 289)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0633-240300-304-g12&t3=2000-03-24&t4=1)
300-304-g12&t3=2000-03-24&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 289.
Location/Qualifiers
1. .289
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0633"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
62 a 74 c 93 g 60 t

FEATURES

source
1. .289
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0633"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
62 a 74 c 93 g 60 t

BASE COUNT

ORIGIN
Query Match 59.7%; Score 123; DB 10; Length 289;
Best Local Similarity 80.4%; Pred. No. 8.5e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTCTGCTTAAAGTCTGCTTAAAGCCAGGTCTTGTCTTATATCT 87
Db 276 AAGCCAGAGCCAGTCAAGGTCCAGTCTCACTAAGCTGGCTCTGCCCATATCT 217
Qy 88 TGATTGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGCCAG 147
Db 216 TGATCCGTTGGCCATGTTGAATCCCTTAACCGTCTGTTGAAAGATCACTGACTGCCAG 157
Qy 148 GTATCAAAAGTGTGTAAGTCTTCGCGGTATGGCTGTTGTTCCCAATAATAG 206
Db 156 GAATCAAGAGTGTGTAAGGCTCTTCGCGGATGGCTGTTCTGTTCCCGCAGTGAGAG 98

RESULT 58
BE181959

LOCUS BE181959 312 bp mRNA linear EST 22-JUN-2000
 DEFINITION CM1-HT0641-280400-206-d02 HT0641 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE181959
 VERSION BE181959.1 GI:8661135
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 312)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM1-HT0641-280400-206-d02&t3=2000-04-28&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 15
 High quality sequence stop: 312.
 FEATURES
 source
 1..312
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0641"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18, Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 70 a 93 c 81 g 68 t
 ORIGIN
 Query Match 59.7%; Score 123; DB 10; Length 312;
 Best Local Similarity 80.4%; Pred. No. 8.7e-29;
 Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCT 87
 DB 69 AAGCGCAAGAGCGGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 128
 QY 88 TGATTCTGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAG 147
 DB 129 TGATCCGGTGGCGCATGTTGAATCCCGCTTAACCGCTGCTTGAAGATAGTACTGCCAG 188
 QY 148 GTATCAAAAGTCTGTGAAGGTTCTGCGGTATGCTTGTTCCTCCACATAATAG 206
 DB 189 GAATCAAGAAGTCTGTGAAGGCTTTTGGGATGGCTTGTTCGTTCCCAAGTGAAG 247
 RESULT 59
 BE15619/c
 LOCUS BE15619 342 bp mRNA linear EST 12-SEP-2000
 DEFINITION CM4-HT0744-160600-201-f01 HT0744 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE15619

VERSION BE15619.1 GI:10103884
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 342)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160600-201-f01&t3=2000-06-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 15
 High quality sequence stop: 342.
 FEATURES
 source
 1..342
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0744"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 78 a 87 c 93 g 83 t
 ORIGIN
 Query Match 59.7%; Score 123; DB 12; Length 342;
 Best Local Similarity 80.4%; Pred. No. 9e-29;
 Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTATTATCT 87
 DB 236 AAGCGCAAGAGCGGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 177
 QY 88 TGATTCTGTTGGCTATGTTAAACCCACCTAACCGTTCTTGAAGGACACTGATTGTCAG 147
 DB 176 TGATCCGGTGGCGCACGTTGAATCCCGCTTAACCGCTGCTTGAAGATAGTACTGCCAG 117
 QY 148 GTATCAAAAGTCTGTGAAGGTTCTGCGGTATGCTTGTTCCTCCCAATAATAG 206
 DB 116 GAATCAAGAAGTCTGTGAAGGCTTTTGGGATGGCTTGTTCGTTCCCAAGTGAAG 58
 RESULT 60
 BF094327/c
 LOCUS BF094327 345 bp mRNA linear EST 19-OCT-2000
 DEFINITION CM4-UT0043-090900-309-b10 UT0043 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF094327
 VERSION BF094327.1 GI:10900037
 KEYWORDS EST.
 SOURCE human.

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 345)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&kt2=QV3-Ht1016-
900-309-b10&kt3=2000-11-09&kt4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 345.
FEATURES
source
1. .345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0043"
/dev_stage="Adult"
/note="Organ: uterus tumor; Vector: puc18; Site:1: SmaI;
Site2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 78 a 86 c 98 g 83 t
ORIGIN
Query Match 59.7%; Score 123; DB 12; Length 345;
Best Local Similarity 80.4%; Pred. No. 9e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 28 ATGCTCAGAACCCAGTTAAAGGTCCTGTCTACTTAAGCCAGGTTCTTGCTTATATCT 87
Db 237 AAGCGCAAGACCCAGTCAAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 178
Qy 88 TGATTCGTTGGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 177 TGATCCGTTGGCCCATGTTGTAATCCCTACCCCTGCTGTAAGATACACTGACTGCCAG 118
Qy 148 GTATCAAAAAGTGTGTAAGGTTCTCGCGTATGGCTTGTTCGTTCCCAATAATAG 206
Db 117 GAATCAAGAAGTGTGTGAAGGCTTTGCGGGATGGCCCGTTTCGTTCCCGAGTGAGAG 59
RESULT 61
BF836201
LOCUS QV3-Ht1016-171100-474-c05 Ht1016 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF836201
ACCESSION BF836201
VERSION BF836201.1 GI:12187891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 374)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&kt2=QV3-Ht1016-
171100-474-c05&kt3=2000-11-17&kt4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 374.
FEATURES
source
1. .374
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site:1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 84 a 98 c 102 g 90 t
ORIGIN
Query Match 59.7%; Score 123; DB 12; Length 374;
Best Local Similarity 80.4%; Pred. No. 9.3e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 28 ATGCTCAGAACCCAGTTAAAGGTCCTGTCTACTTAAGCCAGGTTCTTGCTTATATCT 87
Db 167 AAGCGCAAGACCCAGTCAAAAGTCCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 226
Qy 88 TGATTCGTTGGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 227 TGATCCGTTGGCCCATGTTGTAATCCCTACCCCTGCTGTAAGATACACTGACTGCCAG 286
Qy 148 GTATCAAAAAGTGTGTAAGGTTCTCGCGTATGGCTTGTTCGTTCCCAATAATAG 206
Db 287 GAATCAAGAAGTGTGTGAAGGCTTTGCGGGATGGCCCGTTTCGTTCCCGAGTGAGAG 345
RESULT 62
BI021805/c
LOCUS CM3-MT0316-170101-673-e02 MT0316 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI021805
ACCESSION BI021805
VERSION BI021805.1 GI:14428435
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

```

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-MT0316-170101-673-e02&t3=2001-01-17&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 394.

Location/Qualifiers

1..394

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="MT0316"

/dev_stage="Adult"

/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 91 a 106 c 108 g 89 t

ORIGIN

Query Match 59.7%; Score 123; DB 13; Length 394;

Best Local Similarity 80.4%; Pred. No. 9.5e-29;

Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTTATATCT 87

Db 231 AAGCGAAGAGGAGTCAAGAGTCCAGTCTCCACTAAGCGCTGGCTCTCCCATATCT 172

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAAGCACACTGATGTCAG 147

Db 171 TGATCCGGTGGCCATGTTGAATCCCGCTTAACCGCTGCTTGAAGACATACTGACTGCCAG 112

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAATAATAG 206

Db 111 GAATCAAGAAGTCTGTGAAGGTTCTTGGGATGCGCTTGTTCCTCCAGTGAGAG 53

RESULT 63

AI924155/c

LOCUS

DEFINITION

wn64c08.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2450222 3'

similar to gb:S58717.rna1 ELAFIN PRECURSOR (HUMAN); contains element PTR5 repetitive element ;, mRNA sequence.

AI924155

AI924155.1 GI:5660119

EST.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 420)

REFERENCE

AUTHORS

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 511 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..420

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2450222"

/clone_lib="NCI CGAP Lul9"

/tissue_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"

/dev_stage="adult"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 89 c 130 g 99 t

ORIGIN

Query Match 59.7%; Score 123; DB 9; Length 420;

Best Local Similarity 80.4%; Pred. No. 9.7e-29;

Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTTATATCT 87

Db 360 AAGCGAAGAGGAGTCAAGAGTCCAGTCTCCACTAAGCGCTGGCTCTGCCCATATCT 301

QY 88 TGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAAGCACACTGATGTCAG 147

Db 300 TGATCCGGTGGCCATGTTGAATCCCGCTTAACCGCTGCTTGAAGACATACTGACTGCCAG 241

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGCTTGTTCGTTCCACAATAATAG 206

Db 240 GAATCAAGAAGTCTGTGAAGGTTCTTGGGATGCGCTTGTTCGTTCCAGTGAGAG 182

RESULT 64

BI061069

LOCUS

DEFINITION

IL3-UT0116-020201-464-F04 UT0116 Homo sapiens cDNA, mRNA sequence.

BI061069

ACCESSION

VERSION

EST.

KEYWORDS

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 421)

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE
COMMENT

20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0116-
020201-464-F04&t3=2001-02-02&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 353.

FEATURES

source
1. 421
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0116"
/dev_stage="Adult"

/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 94 a 116 c 113 g 98 t

ORIGIN

Query Match 59.7%; Score 123; DB 13; Length 421;
Best Local Similarity 80.4%; Pred. No. 9.7e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGCTACTTAAGCCAGGTTCTTGCTTATATCT 87
|||
Db 183 AAGCCAGAGACGCTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCT 242
|||
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
|||
Db 243 TGATCCGTTGCCCATGTTGAATCCCCCTAACCGTCTGTTGAAGATAGTACTGACTGCCAG 302
|||
QY 148 GTATCAAAAAGTCTGTGAAGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
|||
Db 303 GAATCAGGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGCAGTGAGAG 361
|||

RESULT 65
AI830232/c

LOCUS AI830232 445 bp mRNA linear EST 21-DEC-1999
DEFINITION wj78g12.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2408998 3'
similar to gb:S58717.fna1 ELAFIN PRECURSOR (HUMAN); contains element
HGR repetitive element ;, mRNA sequence.

ACCESSION AI830232
VERSION AI830232.1 GI:5450892
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 445)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

REFERENCE

AUTHORS Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 531 Std Error: 0.00
Seq primer: -400P from Gibco.

FEATURES

source
1. 445
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2408998"
/clone_lib="NCI-CGAP_Lul9"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"

/dev_stage="adult"

/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p773 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 105 a 97 c 138 g 105 t

ORIGIN

Query Match 59.7%; Score 123; DB 9; Length 445;
Best Local Similarity 80.4%; Pred. No. 9.9e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGCTACTTAAGCCAGGTTCTTGCTTATATCT 87
|||
Db 363 AAGCCAGAGACGCTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCT 304
|||
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
|||
Db 303 TGATCCGTTGCCCATGTTGAATCCCCCTAACCGTCTGTTGAAGATAGTACTGTCGCCAG 244
|||
QY 148 GTATCAAAAAGTCTGTGAAGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
|||
Db 243 GAATCAGGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGCAGTGAGAG 185
|||

RESULT 66

BF835679 447 bp mRNA linear EST 13-JAN-2001
LOCUS QV3-H1016-151100-463-cl0 HT1016 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF835679
ACCESSION BF835679
VERSION BF835679.1 GI:12186947
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 447)
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

REFERENCE

AUTHORS Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922

TITLE

JOURNAL
MEDLINE
COMMENT

Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-HT1016-151100-463-cl06t3=2000-11-15&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 447.
 Location/Qualifiers

1. 447
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT1016"
 /dev_stage="Adult"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 96 a 128 c 117 g 106 t

FEATURES

source

Query Match 59.7%; Score 123; DB 12; Length 447;
 Best Local Similarity 80.4%; Pred. No. 9.9e-29;
 Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGCTCTATTATCT 87
 Db 187 AAGCGGAGAGCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 246
 Qy 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
 Db 247 TGATTCGTTGGCGCATGTTGAATCCCTTACCGCTGCTTGAAGATCTAGTCTGCCAG 306
 Qy 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCGTTCCACAATAATAG 206
 Db 307 GAATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCTTTCGTTCCCGAGTGAGAG 365

RESULT 67
 AW845140/c
 LOCUS AW845140 453 bp mRNA linear EST 19-MAY-2000
 DEFINITION IL0-CT0008-140599-007 CT0008 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW845140
 VERSION AW845140.1 GI:7940657
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 453)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0-CT0008-140599-007&t3=1999-05-14&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 39
 High quality sequence stop: 454.
 Location/Qualifiers

1. 453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0008"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 ,716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 106 a 110 c 139 g 98 t

Query Match 59.7%; Score 123; DB 10; Length 453;
 Best Local Similarity 80.4%; Pred. No. 9.9e-29;
 Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGCTCTATTATCT 87
 Db 309 AAGCGAAGAGCCAGTCAAGGTCAGCTCCACTAAGCCTGGCTCTGCCCATTTATCT 250
 Qy 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
 Db 249 TGATTCGTTGGCGCATGTTGAATCCCTTAAACCGCTGCTTGAAGATCTAGTCTGCCAG 190
 Qy 148 GTATCAAAAAGTCTGTGAAGGTTCTCGCGTATGCTTGTTCGTTCCACAATAATAG 206
 Db 189 GAATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCTTTCGTTCCCGAGTGAGAG 131

RESULT 68
 AW845142/c
 LOCUS AW845142 455 bp mRNA linear EST 19-MAY-2000
 DEFINITION IL0-CT0008-140599-001 CT0008 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW845142
 VERSION AW845142.1 GI:7940659
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 455)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0-CT0008-140599-001&t3=1999-05-14&t4=1)

```

Seq primer: puc 18 forward
High quality sequence stop: 455.
Location/Qualifiers
1. .455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="Adult"
/dev_stage="Adult"
/Note="Organ: colon; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 108 a 116 c 134 g 97 t
ORIGIN

Query Match 59.7%; Score 123; DB 10; Length 455;
Best Local Similarity 80.4%; Pred. No. 9.9e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTTATATCT 87
DB 284 AAGCGCAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCGCTGCTGCCCCATATCT 225

QY 88 TGATTCGTTGGCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATTGTCAG 147
DB 224 TGATCCGTTGGCCATGTTGAATCCCCCTAACCGTCTGTTGAAGATACAGTGCACCAG 165

QY 148 GTATCAAAAGTGTGTAAGGTTCTCGCGTATGGCTGTTTCGTTCCACAAATATAG 206
DB 164 GAATCAAGAAGTGTGTAAGGCTCTTGCAGGATGGCTGTTTCGTTCCCGGTGAGAG 106

RESULT 69
BF824930
LOCUS
DEFINITION ILO-HN0039-101100-500-a05 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824930
VERSION BF824930.1 GI:12166995
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 466)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ILO&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 466.
Location/Qualifiers
1. .466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="Adult"

```

/note="Organ: head normal; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 98 a 138 c 117 g 113 t
 ORIGIN
 Query Match 59.7%; Score 123; DB 12; Length 466;
 Best Local Similarity 80.4%; Pred. No. 1e-28;
 Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGGCTCTGTGTCCTAAAGCCAGGTTCTTGTCTATTATCT 87
 Db 182 AAGCGAAGAGCCAGTCAAGGGTCCAGTCTCCACTAAGCCGTGCTTGAAGATACCTGACTGCCAG 241
 QY 88 TGATTCTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACATGATGTCAG 147
 Db 242 TGATCCGGTCCGCGATGTTGAATCCCCCTAACCGCTGCTTGAAGATACCTGACTGCCAG 301
 QY 148 GTATCAAAAAGTGTGTTGAAGGTTCTTGGCGGTATGCTTGTTCGTTCCACAATAATAG 206
 Db 302 GAATCAAGAAGTGTGTTGAAGGTTCTTGGCGGTATGCTTGTTCGTTCCAGTGAGAG 360

RESULT 71
 BF824920
 LOCUS
 DEFINITION IL0-HN0039-101100-500-c05 HN0039 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF824920
 VERSION BF824920.1 GI:12166980
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 467)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.J.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-HN0039-101100-500-c05&t3=2000-11-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 467.

FEATURES
 source
 1..467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 99 a 137 c 117 g 114 t
 ORIGIN

Query Match 59.7%; Score 123; DB 12; Length 467;
 Best Local Similarity 80.4%; Pred. No. 1e-28;
 Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
 QY 28 ATGCTCAAGAACAGTTAAAGGCTCTGTGTCCTAAAGCCAGGTTCTTGTCTATTATCT 87
 Db 182 AAGCGAAGAGCCAGTCAAGGGTCCAGTCTCCACTAAGCCGTGCTTGGCGGTATGCTTGTTCGTTCCAGTATCT 241
 QY 88 TGATTCTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACATGATGTCAG 147
 Db 242 TGATCCGGTCCGCGATGTTGAATCCCCCTAACCGCTGCTTGAAGATACCTGACTGCCAG 301
 QY 148 GTATCAAAAAGTGTGTTGAAGGTTCTTGGCGGTATGCTTGTTCGTTCCACAATAATAG 206
 Db 302 GAATCAAGAAGTGTGTTGAAGGTTCTTGGCGGTATGCTTGTTCGTTCCAGTGAGAG 360

RESULT 72
 BF825204
 LOCUS
 DEFINITION IL0-HN0039-131100-501-a07 HN0039 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF825204
 VERSION BF825204.1 GI:12167445
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 487)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.J.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-HN0039-131100-501-a07&t3=2000-11-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 11
 High quality sequence stop: 487.
 Location/Qualifiers
 1..487
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
 source
 1..487
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HN0039"
 /dev_stage="Adult"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

BASE COUNT      100 a      146 c      120 g      121 t
ORIGIN

Query Match      59.7%; Score 123; DB 12; Length 487;
Best Local Similarity 80.4%; Pred. No. 1e-28;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCCTATTATCT 87
    |||
Db 203 AAGCGAAGACCAAGTCAAGGTCAGTCTCCACTAAGCCCTGCTCTGCCCAATTATCT 262
    |||
QY 88 TGATTCGTGGCTATGTTAAACCCACTAACCCGTTGTTTGAAGGACACTGATTGTCAG 147
    |||
Db 263 TGATCCGTGGCCATGCTGAATCCCCCTAAACCGCTCTTGAAGATATCTGACTGCCAG 322
    |||
QY 148 GTATCAAAAGTCTGTGAAGGTTCTCGGTATGGCTGTTTCGTTCCACATATAG 206
    |||
Db 323 GAATCAAGAGTCTGTGAAGGCTCTTGGCGGATGGCTCTGTTTCGTTCCCGCAGTGAAG 381
    |||

RESULT 73
AW860590      344 bp      mRNA      linear      EST 19-MAY-2000
LOCUS      QV0-CT0383-030300-139-d12 CT0383 Homo sapiens cDNA, mRNA sequence.
DEFINITION      AW860590
ACCESSION      AW860590.1 GI:7956283
VERSION
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,A.S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,W.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV0-CT0383-030
300-139-d12&t3=2000-03-03&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 54
High quality sequence stop: 344.
Location/Qualifiers
1. .344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0383"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site:1: SmaI; Site:2:
SmaI; A mini-library was made by cloning products derived
from ORSTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      75 a      102 c      82 g      85 t
ORIGIN

```

```

Query Match      59.1%; Score 121.8; DB 10; Length 344;
Best Local Similarity 77.8%; Pred. No. 2.2e-28;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 18 ATACCTGCATATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGTTCTTGT 77
    |||
Db 73 ATAAAGTCTAAAGGCGAAGACCCAGTCTAAGGTCAGTCTCCACTAAGCCCTGGCTCTGC 132
    |||
QY 78 CCTATTATCTTATTGTTGGCTATGTTAAACCCACTAACCCGTTCTTTGAAGGACACT 137
    |||
Db 133 CCCATTATCTTATCGGTCGGTGGCCATGTTGAATCCCTAACCCGCTGCTTGAAGATACT 192
    |||
QY 138 GATTGTCCAGGTATCAAAAAGTGTCTGAAGGTTCTCGGCTATGCGTATGGTGTGTTTCCA 197
    |||
Db 193 GACTGCCCCAGGAATCAAGAGAGCTGTGAAGGCTCTTGGCGGATGGCTGTTTGTCTCCC 252
    |||
QY 198 CAATTAATAG 206
    |||
Db 253 CAGTGAGAG 261
    |||

RESULT 74
AA586983/c
LOCUS      AA586983      386 bp      mRNA      linear      EST 26-SEP-1997
DEFINITION      m969e10.s1 NCI-CGAP Lari1 Homo sapiens cDNA clone IMAGE:1089162 3',
similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ;, mRNA sequence.
ACCESSION      AA586983
VERSION      AA586983.1 GI:2397797
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 386)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 642 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 241.
Location/Qualifiers
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1089162"
/clone_lib="NCI-CGAP Lari1"
/tissue_type="larynx"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: larynx; Vector: Bluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Larynx. 5' adaptor sequence: 5' GAATTCGCACGAG
3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
Average insert size: 0.9 kb."
BASE COUNT      86 a      75 c      121 g      104 t
ORIGIN

Query Match      59.1%; Score 121.8; DB 9; Length 386;
Best Local Similarity 81.5%; Pred. No. 2.3e-28;
Matches 141; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 34 AAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTTATTATTGATTC 93

```


Db 386 AAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTGATCC 327
QY 94 GTTGGCTATGTTAAACCCACCTAACCGTGTGTTCAAGGACACTGATGTCACAGTATCA 153
Db 326 GGTGGCCCATGTGTAATCCCCCTAACCGTGTGTTGAAGAGACTACTGCTGCCAGGAATCA 267
QY 154 AAAAGTGTGTGAAGGTCCTCGGCTATGCTGTTGTTTCGTTCCACAATAATAG 206
Db 266 AGAAGTGTGTGAAGGTCCTCGGCTATGCTGTTTCGTTCCACAATAATAG 214

RESULT 75

BF736734/c

LOCUS

PM1-KT0004-241100-001-c04 KT0004 Homo sapiens cDNA, mRNA sequence.

DEFINITION

BF736734

ACCESSION

BF736734.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

CONTACT

LABORATORY

INSTITUTE

FOR

RESEARCH

RUA

PROF.

ANTONIO

PRUDENTE

109, 4

ANDAR,

01509-010,

SAO PAULO-SP,

BRAZIL

TEL:

+55-11-2704922

FAX:

+55-11-2707001

EMAIL:

asimpson@ludwig.org.br

THIS

SEQUENCE

WAS

DERIVED

FROM

THE

FAPESP/LICR

HUMAN

CANCER

GENOME

PROJECT.

THIS

ENTRY

CAN

BE

SEEN

IN

THE

FOLLOWING

URL

(HTTP://WWW.LUDWIG.ORG.BR/SCRIPTS/GETHTML2.PL?T1=PM1&T2=PM1-KT0004-241100-001-c04&T3=2000-11-24&T4=1)

SEQ

PRIMER:

puc 18 forward

HIGH

QUALITY

SEQUENCE

STOP: 419.

LOCATION/QUALIFIERS

1. .419

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="KT0004"

/dev_stage="Adult"

/note="Organ: bladder tumor; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE

COUNT

99 a

103 c

126 g

91 t

ORIGIN

Query Match

58.9%;

Score

121.4;

DB

12;

Length

419;

Best

Local

Similarity

79.9%;

Pred.

No.

3.2e-28;

Matches

143;

Conservative

0;

Mismatches

36;

Indels

0;

Gaps

0;

QY

28

ATGCTCAAGAACCAAGTCTGCTGCTACTAAGCCAGGCTCTGCTCTATTTATCT

87

Db

287

AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGCTCTGCCCATTTACCT

228

QY

88

TGATTCTGCGGTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTGTCAG

147

Search completed: February 15, 2003, 23:10:13
Job time : 1193 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 23:10:43 ; Search time 67 Seconds
(without alignments)
1565.980 Million cell updates/sec

Title: 09-833799-13A

Perfect score: 206

Sequence: 1 AATTCGAGTCGCTACCATTA.....GTTTCGTTCCACATATAG 206

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications_NA.*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US05_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	125.2	60.8	480	10	US-09-920-300A-1393
c 2	125.2	60.8	480	12	US-10-033-528-1393
c 3	123.2	59.8	2309	10	US-09-954-456-1126
c 4	123.2	59.8	2309	10	US-09-954-456-1126
c 5	32.2	15.6	546	10	US-09-815-242-4199
c 6	32.2	15.6	549	10	US-09-815-242-8354
c 7	30.2	14.7	1489	10	US-09-731-872-213
c 8	30.2	14.7	36159	12	US-10-135-687-3
c 9	29.4	14.3	5150	10	US-09-800-528-14
c 10	29.4	14.1	549	10	US-09-815-242-8688
c 11	29.4	14.1	549	10	US-09-815-242-8956
c 12	29.4	14.1	549	10	US-09-925-637-9
c 13	28.6	13.9	180216	10	US-09-835-232-6
c 14	28.4	13.8	464	10	US-09-864-761-2155
c 15	28.4	13.8	1656	10	US-09-729-674-19
c 16	28.4	13.8	1668	10	US-09-765-205-25
c 17	28.4	13.8	99916	10	US-09-816-095-3
c 18	28.4	13.6	2125	12	US-10-044-090-335
c 19	27.4	13.3	980	10	US-09-785-738A-3

20	27	13.1	354	10	US-09-852-659A-121	Sequence 121, Appl
21	27	13.1	753	9	US-09-852-797-34	Sequence 34, Appl
22	27	13.1	753	10	US-09-853-161-34	Sequence 34, Appl
23	27	13.1	753	10	US-09-852-659A-34	Sequence 34, Appl
24	27	13.1	783	9	US-09-852-797-50	Sequence 50, Appl
25	27	13.1	783	10	US-09-853-161-50	Sequence 50, Appl
26	27	13.1	783	10	US-09-852-659A-50	Sequence 50, Appl
c 27	27	13.1	1536	10	US-09-815-242-6640	Sequence 6640, Ap
28	26.8	13.0	3976	9	US-10-227-884-59	Sequence 59, Appl
29	26.8	13.0	4531	9	US-10-114-893-118	Sequence 118, App
30	26.8	13.0	5686	10	US-09-775-938A-31	Sequence 31, Appl
c 31	26.8	13.0	8210	10	US-09-962-436-279	Sequence 279, App
c 32	26.6	12.9	687	10	US-09-070-927A-612	Sequence 612, App
c 33	26.6	12.9	891	10	US-09-815-242-7107	Sequence 7107, Ap
c 34	26.6	12.9	1857	10	US-09-974-300-6054	Sequence 6054, Ap
35	26.6	12.9	24004	10	US-09-070-927A-72	Sequence 72, Appl
36	26.4	12.8	8921	10	US-09-070-927A-68	Sequence 68, Appl
c 37	26.2	12.7	670	10	US-09-764-869-355	Sequence 355, App
38	26.2	12.7	1784	9	US-10-105-930-1	Sequence 1, Appll
c 39	26.2	12.7	2000	9	US-09-938-842A-4875	Sequence 4875, Ap
c 40	26.2	12.7	3479	10	US-09-822-849A-298	Sequence 298, App
c 41	26.2	12.7	157875	9	US-09-935-464-1	Sequence 1, Appli
c 42	26.2	12.7	640681	10	US-09-790-988-1	Sequence 1, Appli
c 43	26	12.6	1340	9	US-10-214-188-3	Sequence 3, Appli
c 44	26	12.6	2000	9	US-09-938-842A-4030	Sequence 4030, Ap
45	25.8	12.5	340	10	US-09-764-877-2223	Sequence 223, Ap
46	25.8	12.5	539	10	US-09-070-927A-722	Sequence 722, App
47	25.8	12.5	1047	10	US-09-886-055-410	Sequence 410, App
c 48	25.8	12.5	1049	10	US-09-800-729-67	Sequence 67, Appl
c 49	25.8	12.5	1107	9	US-09-738-526-1385	Sequence 1385, Ap
50	25.8	12.5	1125	9	US-09-965-528-31	Sequence 31, Appl
c 51	25.8	12.5	1600	10	US-09-938-540-1	Sequence 1, Appl
c 52	25.8	12.5	3098	10	US-09-925-300-693	Sequence 693, App
c 53	25.8	12.5	3189	12	US-10-044-090-333	Sequence 333, App
c 54	25.8	12.5	3299	10	US-09-800-729-68	Sequence 68, Appl
c 55	25.8	12.5	4116	9	US-09-822-846-104	Sequence 104, App
c 56	25.8	12.5	6065	10	US-09-800-729-35	Sequence 35, Appl
57	25.6	12.4	247	9	US-10-002-048A-4	Sequence 4, Appli
58	25.6	12.4	327	10	US-09-960-352-6148	Sequence 6148, Ap
c 59	25.6	12.4	434	10	US-09-954-456-2208	Sequence 2208, Ap
c 60	25.6	12.4	434	10	US-09-880-107-2920	Sequence 2920, Ap
c 61	25.6	12.4	468	10	US-09-783-590-11050	Sequence 11050, A
c 62	25.6	12.4	633	9	US-09-995-898A-29	Sequence 29, Appl
c 63	25.6	12.4	729	10	US-09-910-943-715	Sequence 715, App
64	25.6	12.4	942	10	US-09-815-242-8813	Sequence 8813, Ap
65	25.6	12.4	942	10	US-09-815-242-8969	Sequence 8969, Ap
66	25.6	12.4	1473	9	US-09-995-898A-3	Sequence 3, Appli
67	25.6	12.4	1560	9	US-09-995-898A-28	Sequence 28, Appl
68	25.6	12.4	2211	10	US-09-764-864-359	Sequence 359, App
69	25.6	12.4	6134	10	US-09-764-877-3301	Sequence 3301, Ap
70	25.6	12.4	28690	10	US-09-070-927A-138	Sequence 138, App
c 71	25.4	12.3	326	10	US-09-998-598-2252	Sequence 2252, Ap
c 72	25.4	12.3	327	9	US-09-796-692-7381	Sequence 7381, Ap
c 73	25.4	12.3	504	10	US-09-783-590-7280	Sequence 7280, Ap
74	25.4	12.3	1026	10	US-09-822-830A-205	Sequence 205, App
75	25.4	12.3	1041	8	US-08-927-839-55	Sequence 55, Appl
76	25.4	12.3	1057	9	US-09-974-298-14	Sequence 14, Appl
77	25.4	12.3	1057	12	US-10-044-090-574	Sequence 574, App
78	25.4	12.3	1103	12	US-10-121-119-6	Sequence 6, Appli
79	25.4	12.3	1104	12	US-10-044-090-575	Sequence 575, App
c 80	25.4	12.3	2084	9	US-10-174-590-175	Sequence 175, App
c 81	25.4	12.3	2084	9	US-10-175-758-175	Sequence 175, App
c 82	25.4	12.3	2084	9	US-10-175-737-175	Sequence 175, App
c 83	25.4	12.3	2084	9	US-10-173-706-175	Sequence 175, App
c 84	25.4	12.3	2084	9	US-10-175-738-175	Sequence 175, App
c 85	25.4	12.3	2084	9	US-10-175-752-175	Sequence 175, App
c 86	25.4	12.3	2084	9	US-10-176-482-175	Sequence 175, App
c 87	25.4	12.3	2084	9	US-10-176-757-175	Sequence 175, App
c 88	25.4	12.3	2084	9	US-10-176-913-175	Sequence 175, App
c 89	25.4	12.3	2084	9	US-10-180-552-175	Sequence 175, App
c 90	25.4	12.3	2084	9	US-10-180-557-175	Sequence 175, App
c 91	25.4	12.3	2084	9	US-10-173-700-175	Sequence 175, App
c 92	25.4	12.3	2084	9	US-10-174-572-175	Sequence 175, App

C 93 25.4 12.3 2084 9 US-10-174-579-175 Sequence 175, App
C 94 25.4 12.3 2084 9 US-10-174-582-175 Sequence 175, App
C 95 25.4 12.3 2084 9 US-10-174-588-175 Sequence 175, App
C 96 25.4 12.3 2084 9 US-10-175-739-175 Sequence 175, App
C 97 25.4 12.3 2084 9 US-10-175-740-175 Sequence 175, App
C 98 25.4 12.3 2084 9 US-10-175-743-175 Sequence 175, App
C 99 25.4 12.3 2084 9 US-10-176-488-175 Sequence 175, App
C 100 25.4 12.3 2084 9 US-10-176-492-175 Sequence 175, App

ALIGNMENTS

RESULT 1

US-09-920-300A-1393/c
; Sequence 1393, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Gordon E.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1393
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 27, 56, 343
; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-1393

Query Match 60.8%; Score 125.2; DB 10; Length 480;
Best Local Similarity 81.0%; Pred. No. 2.8e-34;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 371 AAGCGAAGAGCCAGTCAAAAGTCCAGTNTCCACTAAGCGCTGGCTCTGCCCATTAATCT 312
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTCACAG 147
Db 311 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGTGAAAGATACCTGACTGCCACAG 252
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTTGGCGGTATGCTTGTTCGTTCCACAATAATAG 206
Db 251 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGCGCTTCTTCCCGAGTGGAGAG 193

RESULT 2

US-10-033-528-1393/c
; Sequence 1393, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1393

; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 27, 56, 343
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1393

Query Match 60.8%; Score 125.2; DB 12; Length 480;
Best Local Similarity 81.0%; Pred. No. 2.8e-34;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 371 AAGCGAAGAGCCAGTCAAAAGTCCAGTNTCCACTAAGCGCTGGCTCTGCCCATTAATCT 312
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTCACAG 147
Db 311 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGTGAAAGATACCTGACTGCCACAG 252
QY 148 GTATCAAAAAGTCTGTGAAGTTCCTGCGGTATGCTTGTTCGTTCCACAATAATAG 206
Db 251 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGCGCTTCTTCCCGAGTGGAGAG 193

RESULT 3

US-09-954-456-1126
; Sequence 1126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1126
; LENGTH: 2309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1126

Query Match 59.8%; Score 123.2; DB 10; Length 2309;
Best Local Similarity 81.2%; Pred. No. 2.9e-33;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87

Db 1552 AAGCGCAAGACCCAGTCAAAAGGTCAGTCTCCACTAAGCCCTGGCTCTGCCCCATTATCT 1611
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAAACCGTTCTTTGAAGGACACTGATTGTCCAG 147
Db 1612 TGATCCGGTGGCCATGTTGAATCCCTAACCCCTGCTTGAAGATACACTGACTGCCAG 1671
QY 148 GTATCAAAAAGTGTGTAAGGTTCTGCGGTATGGCTTGTTCGTTCCACAATAA 203
Db 1672 GAATCAAGAAGTGTGTAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCAAGTGA 1727

RESULT 4

US-09-954-456-1798
; Sequence 1798, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cando
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1798
; LENGTH: 2309
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1798

Query Match 59.8%; Score 123.2; DB 10; Length 2309;
Best Local Similarity 81.2%; Pred. No. 2.9e-33;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 1552 AAGCGCAAGACCCAGTCAAAAGGTCAGTCTCCACTAAGCCCTGGCTCTGCCCCATTATCT 1611
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAAACCGTTCTTTGAAGGACACTGATTGTCCAG 147
Db 1612 TGATCCGGTGGCCATGTTGAATCCCTAACCCCTGCTTGAAGATACACTGACTGCCAG 1671
QY 148 GTATCAAAAAGTGTGTAAGGTTCTGCGGTATGGCTTGTTCGTTCCACAATAA 203
Db 1672 GAATCAAGAAGTGTGTAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCAAGTGA 1727

RESULT 5

US-09-815-242-4199/c

; Sequence 4199, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4199
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4199

Query Match 15.6%; Score 32.2; DB 10; Length 546;
Best Local Similarity 48.1%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 98

QY 2 ATTGAGCTCGGTACCATACCTGCATATGCTCAAGAACCTGTTAAAGCTCTGTGTCTAC 61
Db 312 ATTTGGCTTAGACCTGCACCTGCAGAACCTACAAAACCACTAACGCCCTGGTATTCT 253
QY 62 TAAGCCAGGTTCTTGTCTCTATTATCTTGAATTCGTTGGCTATGTTAAACCCACCTAACCG 121
Db 252 TACCACATACCATGATTCATCTGTCATGATTAAATTCACATAAAACATATCCAGGAATGT 193
QY 122 TTGTTTGAAGGACACTGATTCTCAGGTATCAAAAAGTGTGTAAGGTTCTTCGGGTAT 181
Db 192 TTTTAAAGGTTGTTTAGCTTTACCATCTTTAGTTTCTTCTTCTTCGGGTAT 133
QY 182 GGCTTGT 190
Db 132 GACTACTCT 124

RESULT 6

US-09-815-242-8354/c
; Sequence 8354, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: C70237D1
; CURRENT APPLICATION NUMBER: US/09/800,528
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 09/068,140
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: PCT/EP/9604807
; PRIOR FILING DATE: 1996-11-04
; PRIOR APPLICATION NUMBER: GB9522558.7
; PRIOR FILING DATE: 1995-11-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 5150
; TYPE: DNA
; ORGANISM: Ribes nigrum
US-09-800-528-14

Query Match 14.3%; Score 29.4; DB 10; Length 5150;
Best Local Similarity 55.3%; Pred. No. 3.4;
Matches 57; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 100 CTATGTTAAACCCACTAACCGTTGTTTGAAGGACACTGATTGTCAGGTATCAAAAAGT 159
DB 2569 CTAITTAACAAGACTTCTAGTTGTTTCAGGGATTATTATTCATATTTATCAATAAAA 2510
QY 160 GCTGTGAAGTCTCCGCGGTATGCTTGTTCGTTCCACAATA 202
DB 2509 ACTGAGAATTTCTCCGCTGTTCTTCTTCTTACGAGAATA 2467

RESULT 10
US-09-815-242-8688/c
; Sequence 8688, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8688
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(549)

US-09-815-242-8688
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(549)

Query Match 14.1%; Score 29; DB 10; Length 549;
Best Local Similarity 51.1%; Pred. No. 1.6;
Matches 68; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 2 ATTGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGTCTAC 61
DB 312 ATTTGGCTTAGACCTCGACCTGCAGAACCTACAAAACCAAGTACCGCTGGTATTCT 253
QY 62 TAAGCCAGGTTCTTGTCTTATTATCTTGTTCGCTATGTTAAACCCACCTAACCG 121
DB 252 TACCACATACCATGATTCTGTCATGTAATTAATCCACTAAACATATCCAGGAATGT 193
QY 122 TTGTTTGAAGGAC 134
DB 192 TTTTAAACAGTC 180

RESULT 11
US-09-815-242-8956/c
; Sequence 8956, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(549)

US-09-815-242-8956
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(549)

Query Match 14.1%; Score 29; DB 10; Length 549;
Best Local Similarity 51.1%; Pred. No. 1.6;
Matches 68; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 2 ATTGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGTCTAC 61
DB 312 ATTTGGCTTAGACCTCGACCTGCAGAACCTACAAAACCAAGTACCGCTGGTATTCT 253
QY 62 TAAGCCAGGTTCTTGTCTTATTATCTTGAATTCGCTATGTTAAACCCACCTAACCG 121
DB 252 TACCACATACCATGATTCTGTCATGTAATTAATCCACTAAACATATCCAGGAATGT 193
QY 122 TTGTTTGAAGGAC 134

09-833799-13a.rnpb

Sun Feb 16 09:12:56 2003

US-09-835-232-6

Query Match 13.9%; Score 28.6; DB 10; Length 180216;
 Best Local Similarity 53.0%; Pred. No. 35; Gaps 0;
 Matches 61; Conservative 0; Mismatches 54; Indels 0;
 QY 57 TCTACTAGCCAGGTTCTTGTCTTATTTATTTGCTTGGCTATGTTAAACCCACCT 116
 Db 169405 TCTTTTCAGTGATGAATTTCTTAAATTTCTTCACTTAGAATTTTTTTTATCTCACCT 169464
 QY 117 AACCGTTGTTTCAAGGACACACTGATGTCAGGTATCAAAAAGTGTCTGAAGTT 171
 Db 169465 TCTTTTTCAGTGATGAATTTCTTAAATTTGATATAGAGTCTGGGTTGACAGTT 169519

RESULT 14

US-09-864-761-2155
 ; Sequence 2155; Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 2155
 ; LENGTH: 464
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC008014.5

US-09-835-232-6

Query Match 14.1%; Score 29; DB 10; Length 549;
 Best Local Similarity 51.1%; Pred. No. 1.6; Gaps 0;
 Matches 68; Conservative 0; Mismatches 65; Indels 0;
 QY 2 ATTCGAGTCGGTACCATCTGATGTCAGAACACAGTTAAAGGTCCTGTGCTAC 61
 Db 315 ATTTGGCTTAGACCTGACCTGACAGACCTACAAAACACAGTACGGCTGTGTTCT 256
 QY 62 TAAGCCAGTTCTTGTCTTATTTGATTCGTTGGCTATGTTAAACCCACCTAACCG 121
 Db 255 TACCACATACCATGATTCATCTGTCATGATTAATTCACCTAAAAACATATCCAGGGAATG 196
 QY 122 TTGTTTGAAGGAC 134
 Db 195 TTTTAAACAGTC 183

RESULT 13

US-09-835-232-6
 ; Sequence 6; Application US/09835232
 ; Patent No. US20020098489A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leder, Philip
 ; APPLICANT: Leader, Benjamin
 ; TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 00383/052002
 ; CURRENT APPLICATION NUMBER: US/09/835,232
 ; CURRENT FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: US 60/196,811
 ; PRIOR FILING DATE: 2000-04-13
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 180216
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(180216)
 ; OTHER INFORMATION: n = A,T,C or G

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
US-09-864-761-2155

Query Match 13.8%; Score 28.4; DB 10; Length 464;
Best Local Similarity 54.9%; Pred. No. 2.4;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 77 TCCTATTATCTTGATTCGTCGGTATGTTAAACCCACCTAACCGTTGTTGAAGACAC 136
Db 293 TCCTCTTTCCACAGATTATTGCGTCTGCTCCCAAGCCCTTCACTGATAGATGAAGCCCTG 352
QY 137 TGATTGTCCAGGTATCAAAAGTCTGTGAAGGTTCTCTGCGG 178
Db 353 TGATTTTCTGGCATCATGATGACTGAGAGGTCAGTGTGG 394

RESULT 15

US-09-729-674-19/c
; Sequence 19, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fecthel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-19

Query Match 13.8%; Score 28.4; DB 10; Length 1656;
Best Local Similarity 62.9%; Pred. No. 4.4;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 32 TCAAGAACCAAGTAAAGGTCCTGTGCTACTAAAGCCAGGTTCTTGTCCTATTATCTTGAT 91
Db 1023 TCAAGCTACAGGTAAGGCTCTACCAACCAATGAAGTATTGACTTATCCTATGAATTGAT 964
QY 92 TCGTTGCGCT 101
Db 963 TCTTTCTTCT 954

RESULT 16

US-09-765-205-25/c

; Sequence 25, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458.004/200130.449
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 1668
; TYPE: DNA
; ORGANISM: human
US-09-765-205-25

Query Match 13.8%; Score 28.4; DB 10; Length 1668;
Best Local Similarity 62.9%; Pred. No. 4.4;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 32 TCAAGAACCAAGTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGTCCTATTATCTTGAT 91
Db 1094 TCAAGCTACAGTAAAGGCTCTACCAACCAATTAAAGGTTAGTCTTATCTCTATGAATTTGAT 1035
QY 92 TCGTTGCGCT 101
Db 1034 TCTTTCTTCT 1025

RESULT 17

US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match 13.8%; Score 28.4; DB 10; Length 99916;
Best Local Similarity 51.6%; Pred. No. 31;
Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 77 TCCTATTATCTTGATTCGTCGGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACAC 136
Db 75420 TACTATTTTCTACAGCAGCTGCATCATTTTACACTCTCCACAGCAATGCAGGAAGTTCC 75479
QY 137 TGATTGTCAGGTATCAAAAAGTCTGTGAAGGTTCTTCGGGTATGCTTGTTCGTTC 196
Db 75480 AGATTCTCATATCCTCAATAACTATTATTGCTTCTCTTTTGTGTATTAGTTTTTA 75539
QY 197 ACAATA 202
Db 75540 ATAATA 75545

RESULT 18

```

RESULT 20
US-09-852-659A-121
; Sequence 121, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 121
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-121

Query Match      13.1%; Score 27; DB 10; Length 354;
Best Local Similarity 57.8%; Pred.No. 6.5; Indels 0; Gaps
Matches 48; Conservative 0; Mismatches

QY   123 TCGTTGAAGGACACTGATGTGCCAGTGATCAAAAGTGTGTGAAGGTTCCTCGCGGTATG 182
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB    94 TGTAACAAGGACAGACATGCCAGGACACAAAGAAGTGTGTCTTCAGCTCGCGAAAA 153
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY   183 GCTTGTGTTTCGTTCCCAATAATA 205
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB    154 AAATGTTTAGATCTCAAACAAGA 176
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 21
US-09-852-797-34
; Sequence 34, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12

```


Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 24

US-09-852-797-50
; Sequence 50, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-797-50

Query Match 13.1%; Score 27; DB 9; Length 783;

Best Local Similarity 57.8%; Pred. No. 9.5;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 123 TGTTTGAAGGACACTGATTCTCCAGGTATCAAAAAGTGTGTGAAGTTCTCTCGCGGTATG 182

Db 172 TGTACAAGGACACACATCCAGGACACAAAGAAGTTGTCTTTCAGCTCGGAAAA 231

Qy 183 GCTTGTTCGTTCCACAATAA 205

Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 25

US-09-853-161-50
; Sequence 50, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-161-50

Query Match 13.1%; Score 27; DB 10; Length 783;

Best Local Similarity 57.8%; Pred. No. 9.5;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 123 TGTTTGAAGGACACTGATTCTCCAGGTATCAAAAAGTGTGTGAAGTTCTCTCGCGGTATG 182

Db 172 TGTACAAGGACACACATCCAGGACACAAAGAAGTTGTCTTTCAGCTCGGAAAA 231

Qy 183 GCTTGTTCGTTCCACAATAA 205

Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 26

US-09-852-659A-50
; Sequence 50, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368

;; PRIOR FILING DATE: 1997-12-19
;; NUMBER OF SEQ ID NOS: 121
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 50
;; LENGTH: 783
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-852-659A-50

Query Match 13.1%; Score 27; DB 10; Length 783;
Best Local Similarity 57.8%; Pred. No. 9.5;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 123 TGTGTAAGGACACTGATTTCCAGGTATCAAAAAAGTCTGTGAAGGTTCTTCGCGGTATG 182

Db 172 TGTACAAAGGACAGACAATGCCAGGACACAAAGAAGTGTGTCTTCAGCTGCGGAAA 231

QY 183 GCTTGTTCCTCCACAATAA 205

Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 27

US-09-815-242-6640/c

;; Sequence 6640, Application US/09815242

;; Patent No. US20020061569A1

;; GENERAL INFORMATION:

;; APPLICANT: Haselbeck, Robert

;; APPLICANT: Ohlsen, Kari L.

;; APPLICANT: Zyskind, Judith W.

;; APPLICANT: Wall, Daniel

;; APPLICANT: Trawick, John D.

;; APPLICANT: Carr, Grant J.

;; APPLICANT: Yamamoto, Robert T.

;; APPLICANT: Xu, H. Howard

;; TITLE OF INVENTION: Identification of Essential Genes in

;; FILE REFERENCE: ELITRA.011A

;; CURRENT APPLICATION NUMBER: US/09/815,242

;; CURRENT FILING DATE: 2001-03-21

;; PRIOR APPLICATION NUMBER: 60/191,078

;; PRIOR FILING DATE: 2000-03-21

;; PRIOR APPLICATION NUMBER: 60/206,848

;; PRIOR FILING DATE: 2000-05-23

;; PRIOR APPLICATION NUMBER: 60/207,727

;; PRIOR FILING DATE: 2000-05-26

;; PRIOR APPLICATION NUMBER: 60/242,578

;; PRIOR FILING DATE: 2000-10-23

;; PRIOR APPLICATION NUMBER: 60/253,625

;; PRIOR FILING DATE: 2000-11-27

;; PRIOR APPLICATION NUMBER: 60/257,931

;; PRIOR FILING DATE: 2000-12-22

;; PRIOR APPLICATION NUMBER: 60/269,308

;; PRIOR FILING DATE: 2001-02-16

;; NUMBER OF SEQ ID NOS: 14110

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 6640

;; LENGTH: 1536

;; TYPE: DNA

;; ORGANISM: Enterococcus faecalis

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: (1)...(1536)

US-09-815-242-6640

Query Match 13.1%; Score 27; DB 10; Length 1536;

Best Local Similarity 57.8%; Pred. No. 13;

Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 44 TAAAGTCTGTCTACTAAGCAGGTTCTTGCTCTATTATCTTGATTCGTTGCGCTAT 103

Db 1455 TAAAGTCTTTTCTAGCACTGCTGTTCTTCTGATCTCTGATCTATTCGAC 1396

QY 104 GTTAAACCCACCTAACCGTTGTT 126

Db 1395 ATGAACCATATCAATTGGTTCTT 1373

RESULT 28

US-10-227-884-59

;; Sequence 59, Application US/10227884

;; Publication No. US20030027988A1

;; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.

;; APPLICANT: Desnoyers, Luc

;; APPLICANT: Gerritsen, Mary

;; APPLICANT: Goddard, Audrey

;; APPLICANT: Godowski, Paul J.

;; APPLICANT: Grimaldi, J. Christopher

;; APPLICANT: Gurney, Austin L.

;; APPLICANT: Smith, Victoria

;; APPLICANT: Stephan, Jean-Philippe F.

;; APPLICANT: Watanabe, Colin L.

;; APPLICANT: Wood, William I.

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

;; FILE REFERENCE: P3530P1C79

;; CURRENT APPLICATION NUMBER: US/10/227,884

;; CURRENT FILING DATE: 2002-08-26

;; PRIOR APPLICATION NUMBER: 10/119,480

;; PRIOR FILING DATE: 2002-04-09

;; PRIOR APPLICATION NUMBER: 60/059113

;; PRIOR FILING DATE: 1997-09-17

;; PRIOR APPLICATION NUMBER: 60/062287

;; PRIOR FILING DATE: 1997-10-17

;; PRIOR APPLICATION NUMBER: 60/063549

;; PRIOR FILING DATE: 1997-10-28

;; PRIOR APPLICATION NUMBER: 60/064103

;; PRIOR FILING DATE: 1997-10-31

;; PRIOR APPLICATION NUMBER: 60/069873

;; PRIOR FILING DATE: 1997-12-17

;; PRIOR APPLICATION NUMBER: 60/078910

;; PRIOR FILING DATE: 1998-03-20

;; PRIOR APPLICATION NUMBER: 60/079294

;; PRIOR FILING DATE: 1998-03-25

;; PRIOR APPLICATION NUMBER: 60/079656

;; PRIOR FILING DATE: 1998-03-26

;; PRIOR APPLICATION NUMBER: 60/079728

;; PRIOR FILING DATE: 1998-03-27

;; PRIOR APPLICATION NUMBER: 60/081819

;; PRIOR FILING DATE: 1998-04-15

;; PRIOR APPLICATION NUMBER: 60/081955

;; PRIOR FILING DATE: 1998-04-15

;; PRIOR APPLICATION NUMBER: 60/082804

;; PRIOR FILING DATE: 1998-04-22

;; PRIOR APPLICATION NUMBER: 60/084441

;; PRIOR FILING DATE: 1998-05-06

;; PRIOR APPLICATION NUMBER: 60/085323

;; PRIOR FILING DATE: 1998-05-13

;; PRIOR APPLICATION NUMBER: 60/085579

;; PRIOR FILING DATE: 1998-05-15

;; PRIOR APPLICATION NUMBER: 60/086392

;; PRIOR FILING DATE: 1998-05-22

;; PRIOR APPLICATION NUMBER: 60/089532

;; PRIOR FILING DATE: 1998-06-17

;; PRIOR APPLICATION NUMBER: 60/089538

;; PRIOR FILING DATE: 1998-06-17

;; PRIOR APPLICATION NUMBER: 60/089905

;; PRIOR FILING DATE: 1998-06-18

;; PRIOR APPLICATION NUMBER: 60/090472

;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090557

;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090691

;; PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090695

09-833799-13a.rnpb

Sun Feb 16 09:12:56 2003

```

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/095302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

```

Query Match 13.0%; Score 26.8; DB 9; Length 3976;

Best Local Similarity 50.8%; Pred. No. 24;

Matches 64; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```

Qy 37 AACCAAGTTAAAGGTCCTGCTCTACTAAGCAGGTCCTTGTCTCTATTTATCTGTTGTT 96
Db 2933 AATAATTTCAAGTTGAAGTGTTTTTAAAAACACTTTCTTTTGAATGTTTGAATCTCTT 2992
Qy 97 GCGTATGTTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAGGTATCAAAA 156
Db 2993 GAGATGTTGTTTACCCCACTAGATACATATTTGCCACTGGTTAGTTCTCCATCAAGCTCA 3052

```

QY 157 AGTGCT 162
|||
Db 3053 AGAGGT 3058

RESULT 29

US-10-114-893-118
; Sequence 118, Application US/10114893
; Publication No. US20020193567A1

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.

; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

; FILE REFERENCE: GI 6000-10A

; CURRENT APPLICATION NUMBER: US/10/114,893

; CURRENT FILING DATE: 2002-04-02

; EARLIER APPLICATION NUMBER: 09/413,232

; EARLIER FILING DATE: 1999-10-06

; NUMBER OF SEQ ID NOS: 321

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 118

; LENGTH: 4531

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-114-893-118

Query Match

Best Local Similarity 13.0%; Score 26.8; DB 9; Length 4531;

Mismatches 0; Pred. No. 26;

Matches 64; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 37 AACGAGTAAGGTCCTGCTCTACTAAGCCAGGTTCTTGCTCTATATCTGATTCGTT 96

Db 3470 AATAATTTCAAGTTGAAGTGTTTTAAAAACACTTTGTTTGTGAATGTTTGAATCTCT 3529

QY 97 GCGCTATGTTAAACCCACCTAAACGTTGTTGAAGGACACTGATTCGCCAGGTATCAAAA 156

Db 3530 GAGATGTTTACCCCACTAGATACATATTTGCCACTGGTTAGTTCTCCATCTAAGCTCA 3589

QY 157 AGTGCT 162

|||

Db 3590 AGAGGT 3595

RESULT 30

US-09-775-938A-31/c

; Sequence 31, Application US/09775938A

; Patent No. US20020081665A1

; GENERAL INFORMATION:

; APPLICANT: Haygood, M.

; APPLICANT: Davidson, S.K.

; APPLICANT: Allen, S.W.

; APPLICANT: Hildebrand, M.

; TITLE OF INVENTION: Bryostatins, Bryopyrans and Polyketides: Compositions and Methods

; FILE REFERENCE: 1133 010US1

; CURRENT APPLICATION NUMBER: US/09/775,938A

; CURRENT FILING DATE: 2001-08-30

; PRIOR APPLICATION NUMBER: PCT/US00/21326

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: US 60/147,283

; PRIOR FILING DATE: 1999-08-04

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 31

; LENGTH: 5686

; TYPE: DNA

; ORGANISM: Endobugula sertula

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(5686)

; OTHER INFORMATION: N refers to any nucleotide.

US-09-775-938A-31

Query Match

Best Local Similarity 13.0%; Score 26.8; DB 10; Length 5686;

Mismatches 49; Conservative 6; Mismatches 43; Indels 0; Gaps 0;

QY 30 GCTCAAGAACCAAGTTAAAGGTCCTGCTCTACTAAGCCAGGTTCTTGCTCTATATCTGTTG 89

Db 2495 GCCCAAGTTGCAATGCTWTWCGCAARAWCAATTAAAGCSMAWTTCCCTACCTTTTGGCTTC 2436

QY 90 ATTGCTTCGCTATGTTTAAACCCACCTAACCGTTGTTT 127

Db 2435 ATTTTTTYGATAAACTTAAGCAAAACGATCCMYTWTIT 2398

RESULT 31

US-09-962-436-279

; Sequence 279, Application US/09962436

; Patent No. US20020081301A1

; GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat

; FILE REFERENCE: 689290-75

; CURRENT APPLICATION NUMBER: US/09/962,436

; CURRENT FILING DATE: 2001-09-25

; PRIOR APPLICATION NUMBER: US/60/235,082

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/234,924

; PRIOR FILING DATE: 2000-09-25

; NUMBER OF SEQ ID NOS: 568

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 279

; LENGTH: 8210

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-962-436-279

Query Match

Best Local Similarity 13.0%; Score 26.8; DB 10; Length 8210;

Mismatches 58; Conservative 0; Mismatches 52; Indels 0; Gaps 0

QY 33 CAAGAACCAAGTTAAAGGTCCTGCTCTACTAAGCCAGGTTCTTGCTCTATATCTGTTG 92

Db 6952 CATGNACCATTTCTCGNACATTTGGCCCTATATTCAGGGTTTGATGATGTTGGAAATT 7011

QY 93 CGTTGCGCTATGTTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTG 142

Db 7012 CTTGCTTCTTCAGAAACCAAGAGATTCTTTTAAAGGCTCAGATGG 7061

RESULT 32

US-09-070-927A-612/c

; Sequence 612, Application US/09070927A

; Patent No. US20020120116A1

; GENERAL INFORMATION:

; APPLICANT: Charles A. Kunsch

; APPLICANT: Patrick J. Dillon

; APPLICANT: Steven Barash

; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

; NUMBER OF SEQUENCES: 982

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

Sun Feb 16 09:12:56 2003

```

; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 612:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 612:
US-09-070-927A-612

```

```

Query Match 12.9%; Score 26.6; DB 10; Length 687;
Best Local Similarity 54.1%; Pred. No. 12;
Matches 53; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 AATTGAGCTCGGTACCATCTGATATGCTCAAGAACCATGTTAAAGTCTGTGCTA 60
DB 534 AATACCGCTAATAACAATCCCTGATATGCTGCTGGAATATTTTAGCTGCTGCTCAA 475

QY 61 CTAAAGCAGGTTCTGTGCTCTATATCTTGATTCGTGC 98
DB 474 CTACTTCTTGTGCTGACCAACAAGAGATTCGTNCC 437

```

```

RESULT 33
US-09-815-242-7107/c
; Sequence 7107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7107
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(891)
US-09-815-242-7107

```

```

Query Match 12.9%; Score 26.6; DB 10; Length 891;
Best Local Similarity 78.0%; Pred. No. 14;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 154 AAAAGTGTGTAAGGTTCCGCGGTATGGCTTGTTCGTT 194
DB 416 AATTGTGAGTGAAGGTTTGTGCTGTTTGTTCGTT 376

```

```

RESULT 34
US-09-974-300-6054
; Sequence 6054, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6054
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6054

```

```

Query Match 12.9%; Score 26.6; DB 10; Length 1857;
Best Local Similarity 58.0%; Pred. No. 20;
Matches 47; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 25 CATATGCTCAAGAACCATGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCTATTATTA 84
DB 1376 CATTGGCGCTAAACCCCTTTTTCGACCCATTACCAAGAAATTGACGCGGTGGCTGATTC 1435

QY 85 TCTTGATTCGTTGGCTATGT 105
DB 1436 CTTTCATACGTTACGCAATGT 1456

```

```

RESULT 35
US-09-070-927A-72
; Sequence 72, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

```


09-833799-13a.rnpb

Sun Feb 16 09:12:56 2003

FILE REFERENCE: SCRIPI300-3
 CURRENT APPLICATION NUMBER: US/09/938,842A
 CURRENT FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: US 60/227,866
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: US 60/264,647
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/300,111
 PRIOR FILING DATE: 2001-06-22
 NUMBER OF SEQ ID NOS: 5379
 SEQ ID NO 4875
 LENGTH: 2000
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-938-842A-4875

Query Match 12.7%; Score 26.2; DB 9; Length 2000;
 Best Local Similarity 58.2%; Pred. No. 28;
 Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
 QY 31 CTCAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTCTATTTCTTGA 90
 Db 843 CACGAACACATGTAACACTTCTAGTGTCAATAAAATTTATGTTGTTAATATAGTAA 784
 QY 91 TTCGTTGGCTATGTTAAA 109
 Db 783 TCAGATGCAAAAGTTAAAA 765

RESULT 40
 US-09-822-849A-298/c
 Sequence 298, Application US/09822849A
 Patent No. US20020045170A1
 GENERAL INFORMATION:
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fechtel, Kim
 APPLICANT: Agostino, Michael J.
 APPLICANT: Howes, Steven H.
 APPLICANT: Resnick, Richard J.
 APPLICANT: Gulukota, Kamalakara
 APPLICANT: Graham, James R.
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
 FILE REFERENCE: GIN 6403
 CURRENT APPLICATION NUMBER: US/09/822,849A
 CURRENT FILING DATE: 2001-09-04
 PRIOR APPLICATION NUMBER: 60/195,582
 PRIOR FILING DATE: 2000-04-06
 NUMBER OF SEQ ID NOS: 598
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 298
 LENGTH: 3479
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-822-849A-298

Query Match 12.7%; Score 26.2; DB 10; Length 3479;
 Best Local Similarity 50.4%; Pred. No. 37;
 Matches 64; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 QY 1 AATTCGAGTCGGTACCATACCTGTCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTA 60
 Db 1948 AATTCGAGTTGTAAATAATATGTCATATAATTTGTTGCCCATATTTGTTGTTT 1889
 QY 61 CTAAGCCAGGTTCTTGTCTCTATTTATTTGATTCGTTGGCTATGTTAAACCCACCTAAC 120
 Db 1888 AAGAGATGGGGTCTGGCTCTGCTCTGTCGAGCTGGAATGATAGTCATACAGCTCACT 1829
 QY 121 GTTCTTT 127
 Db 1828 GCAGCTT 1822

NAME/KEY: SITE
 LOCATION: (22)
 OTHER INFORMATION: n equals a.t.g, or c
 NAME/KEY: SITE
 LOCATION: (24)
 OTHER INFORMATION: n equals a.t.g, or c
 US-09-764-869-355

Query Match 12.7%; Score 26.2; DB 10; Length 670;
 Best Local Similarity 51.5%; Pred. No. 17;
 Matches 52; Conservative 3; Mismatches 46; Indels 0; Gaps 0;
 QY 90 ATTGTTGGCTATGTTAAACCCACTTAACCGTTCTTGTGAAGGACACTGATGTCAGGT 149
 Db 253 AATCTTAGGTTGGATATAAAGGCCCTTCTTGTCTAAAGAGAGTAATCTCCATT 194
 QY 150 ATCAAAAGTGCTGTGAAGGTTCTTGGCGTATGGCTTGT 190
 Db 193 ACKGTAAAGATTCTAAWAGTTTCATGGCATCGACWATTTT 153

RESULT 38
 US-10-105-930-1
 Sequence 1, Application US/10105930
 Publication No. US20030009018A1
 GENERAL INFORMATION:
 APPLICANT: Maeda, Masatsugu
 APPLICANT: Yaguchi, No. US20030009018A1iko
 TITLE OF INVENTION: NOVEL HEMOPOIETIN RECEPTOR PROTEIN, NR12
 FILE REFERENCE: 06501-105U1
 CURRENT APPLICATION NUMBER: US/10/105,930
 CURRENT FILING DATE: 2002-03-25
 PRIOR APPLICATION NUMBER: PCT/JP00/06654
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: JP 2000-240397
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: JP 11-273358
 PRIOR FILING DATE: 1999-09-27
 NUMBER OF SEQ ID NOS: 77
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 1784
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (98) ... (1108)
 US-10-105-930-1

Query Match 12.7%; Score 26.2; DB 9; Length 1784;
 Best Local Similarity 52.3%; Pred. No. 27;
 Matches 58; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 QY 81 ATTATCTTGATTCGTTGGCTATGTTAAACCCACTTAACCGTTGTTTGAAGGACACTGAT 140
 Db 1321 AGTTTGATGTCGAGTGAAGTATGATCATCCCACTGCTCTCTAGCTGGAGGAAGACCA 1380
 QY 141 TGTCAGGTATCAAAAGTGTGGAAGTTCTTGGCGTATGGCTGTTTC 191
 Db 1381 AGACCCCTGTTTCTTAAAGTTTAAACAGCCAGGTGCAGTGGCTATGTC 1431

RESULT 39
 US-09-938-842A-4875/c
 Sequence 4875, Application US/09938842A
 Patent No. US20020160378A1
 GENERAL INFORMATION:
 APPLICANT: Harper, Jeff
 APPLICANT: Kreps, Joel
 APPLICANT: Wang, Xun
 APPLICANT: Zhu, Tong
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 TITLE OF INVENTION: SAME, AND METHODS OF USE


```

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 722:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 722:
US-09-070-927A-722
Query Match 12.5%; Score 25.8; DB 10; Length 539;
Best Local Similarity 48.9%; Pred. No. 21; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 69;
QY 57 TCTACTAAGCAGGCTTCCTGCTATTATCTTGATTGCTGTGGCGCTATGTTAAACCCACT 116
DB 28 TCTAACATGCTCGTGNCTTTTCTTCAACCTTAGCTAGTAGCTTTAGGINTCTTTTCACAC 87
QY 117 AACCGTGTCTTTGAAGGACACTGATTGCCAGGTATCAAAAAGTCTGTCGAAGTTCTCTGC 176
DB 88 TTCCTTTCCCTAAGTATTTCTTTGTAAACCGTCATGGAAGAAAGTAATCTCGCTTTCAAC 147
QY 177 GGTATGGCTTGTGTTTC 191
DB 148 GTTAAAAATTGGATC 162
RESULT 47
US-09-886-055-410
; Sequence 410, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 410
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-886-055-410
Query Match 12.5%; Score 25.8; DB 10; Length 1047;

```


Sun Feb 16 09:12:56 2003

```
; Sequence 1, Application US/09938540
; Patent No. US20020151001A1
; GENERAL INFORMATION:
; APPLICANT: Degussa AG
; TITLE OF INVENTION: New nucleotide sequences which code for the ccpA1 gene
; FILE REFERENCE: 000059 BT
; CURRENT APPLICATION NUMBER: US/09/938,540
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(1388)
; OTHER INFORMATION: ccpA1-Gen
US-09-938-540-1

Query Match      12.5%; Score 25.8; DB 10; Length 1600;
Best Local Similarity 58.4%; Pred. No. 35;
Matches 45; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 50 TCCGTGCTCTACTAAGCCAGGCTTGTCTCTATTATCTTGGTTCGCTATGTTAA 109
Db 728 TCCCTCTTAAACCAAGTCACTGACTTGTCTCTGACTGATGTGGCCACCATGAT 669

QY 110 CCACCTAACGTTGTT 126
Db 668 TCCATCGAGGGTGGT 652

RESULT 52
US-09-925-300-693/c
; Sequence 693, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 693
; LENGTH: 3098
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-693

Query Match      12.5%; Score 25.8; DB 10; Length 3098;
Best Local Similarity 52.3%; Pred. No. 48;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
Db 2612 TTCTTCCCTCTGATCTGATCATCAGTCTGCTGAGCAGACATATTACTGCTGTGGATA 2553

QY 131 GGACACTGATGTCAGGATCAAAAAGTCTGTGAAGGTTCTCTCGCGT 179
Db 2552 GTAAGACTGCTGTGGGGGCTGAGGAGGGGTATGAAGGCTGTCTGGGT 2504

RESULT 53
US-10-044-090-333/c
; Sequence 333, Application US/10044090
; Patent No. US20020137081A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 333
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2960079CBI
US-10-044-090-333

Query Match      12.5%; Score 25.8; DB 12; Length 3189;
Best Local Similarity 52.3%; Pred. No. 49;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
Db 2743 TTCTTCCCTCTGATCTGATCATCAGTCTGCTGAGCAGACATATTACTGCTGTGGATA 2684

QY 131 GGACACTGATGTCAGGATCAAAAAGTCTGTGAAGGTTCTCTCGCGT 179
Db 2683 GTAAGACTGCTGTGGGGGCTGAGGAGGGGTATGAAGGCTGTCTGGGT 2635

RESULT 54
US-09-800-729-68/c
; Sequence 68, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-729-68

Query Match      12.5%; Score 25.8; DB 10; Length 3299;
Best Local Similarity 52.3%; Pred. No. 50;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
Db 2813 TTCTTCCCTCTGATCTGATCATCAGTCTGCTGAGCAGACATATTACTGCTGTGGATA 2754

QY 131 GGACACTGATGTCAGGATCAAAAAGTCTGTGAAGGTTCTCTCGCGT 179
Db 2753 GTAAGACTGCTGTGGGGGCTGAGGAGGGGTATGAAGGCTGTCTGGGT 2705

RESULT 55
US-09-822-846-104/c
; Sequence 104, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
```

APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fecthel, Kim
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6400
CURRENT APPLICATION NUMBER: US/09/822,846
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,605
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 629
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 104
LENGTH: 4116
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-846-104

Query Match 12.5%; Score 25.8; DB 9; Length 4116;
Best Local Similarity 52.3%; Pred. No. 55;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGCTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
DB 977 TTCTTTCCCTCTGATCTGAATCAGCTGAGCTGTGAGCAGACATATTACTCTCTGGATA 918

QY 131 GGACACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCCCTGCGGT 179
DB 917 GTAAGACTGCTGTGGGGCTGAGGAAGGGTATGAAGGCTGCTGGGT 869

RESULT 56
US-09-800-729-35/c
Sequence 35, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044p1
CURRENT APPLICATION NUMBER: US/09/800,729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 35
LENGTH: 6065
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6035)
OTHER INFORMATION: n equals a.t.g, or c
NAME/KEY: SITE
LOCATION: (6037)
OTHER INFORMATION: n equals a.t.g, or c
NAME/KEY: SITE
LOCATION: (6038)
OTHER INFORMATION: n equals a.t.g, or c

US-09-800-729-35

Query Match 12.5%; Score 25.8; DB 10; Length 6065;
Best Local Similarity 52.3%; Pred. No. 67;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGCTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
DB 2813 TTCTTTCCCTCTGATCTGAATCAGCTGAGCTGTGAGCAGACATATTACTCTGTGGATA 2754

QY 131 GGACACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCCCTGCGGT 179
DB 2753 GTAAGACTGCTGTGGGGCTGAGGAAGGGTATGAAGGCTGCTGGGT 2705

RESULT 57

US-10-002-048A-4
Sequence 4, Application US/10002048A
Publication No. US20020182616A1
GENERAL INFORMATION:
APPLICANT: Wahlestedt, Claes
APPLICANT: Ding, Bo
TITLE OF INVENTION: Single Nucleotide Polymorphisms
FILE REFERENCE: 10806-143
CURRENT APPLICATION NUMBER: US/10/002,048A
CURRENT FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: SE 0004035-2
PRIOR FILING DATE: 2000-11-03
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 247
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: Intron
LOCATION: (1)..(29)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (218)..(247)
OTHER INFORMATION:
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: GenBank/M14296
DATABASE ENTRY DATE: 1995-01-08
US-10-002-048A-4

Query Match 12.4%; Score 25.6; DB 9; Length 247;
Best Local Similarity 55.7%; Pred. No. 17;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 14 TACCATACCTGCATATGCTCAAGAACCCAGTTAAAGTCTCTGTCTACTAAAGCCAGGTTTC 73
DB 16 TTCTGTGCTCGAGATGCTAGGTAAACGAGCTGGGGTGTCCGAGCTGACCTCGCC 75

QY 74 TTGTCCTATTATCTTGATTCGTTGCGCT 101
DB 76 TGTCCCTGCTGTGCTGCTGGGTGCGCT 103

RESULT 58

US-09-960-352-6148
Sequence 6148, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352

09-833799-13a.rnpb

Sun Feb 16 09:12:56 2003

Db 154 AATT 151

RESULT 60

US-09-880-107-2920/c

Sequence 2920, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scheff, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2920

LENGTH: 434

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 R61374

NAME/KEY: unsure

LOCATION: (1)..(434)

OTHER INFORMATION: n = a or c or g or t

US-09-880-107-2920

Query Match 12.4%; Score 25.6; DB 10; Length 434;

Best Local Similarity 62.5%; Pred. No. 22;

Mismatches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTGTCAGGATCAAAAAGTGTGTAAGGTTCTCGCGTATGGCTTGT 190

Db 214 GAACGTGTTGCCAAGGATTTAAAGGGCTTCTCGCTCTCTCTTGTATT 155

QY 191 CGTT 194

Db 154 AATT 151

RESULT 61

US-09-783-590-11050

Sequence 11050, Application US/09783590

Patent No. US20020110850A1

GENERAL INFORMATION:

APPLICANT: Dillion, Patrick J.

APPLICANT: Haseltine, William A.

APPLICANT: Li, Haodong

APPLICANT: Rosen, Steven M.

APPLICANT: Ruben, Steven M.

TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2

FILE REFERENCE: PO-16.2C1

CURRENT APPLICATION NUMBER: US/09/783,590

CURRENT FILING DATE: 2000-02-15

PRIOR APPLICATION NUMBER: 08/420,856

PRIOR FILING DATE: 1995-04-12

PRIOR APPLICATION NUMBER: 08/346,731

PRIOR FILING DATE: 1994-11-21

NUMBER OF SEQ ID NOS: 12485

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 11050

LENGTH: 468

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

Db 154 AATT 151

RESULT 60

US-09-880-107-2920/c

Sequence 2920, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scheff, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2920

LENGTH: 434

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 R61374

NAME/KEY: unsure

LOCATION: (1)..(434)

OTHER INFORMATION: n = a or c or g or t

US-09-880-107-2920

Query Match 12.4%; Score 25.6; DB 10; Length 434;

Best Local Similarity 62.5%; Pred. No. 22;

Mismatches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTGTCAGGATCAAAAAGTGTGTAAGGTTCTCGCGTATGGCTTGT 190

Db 214 GAACGTGTTGCCAAGGATTTAAAGGGCTTCTCGCTCTCTCTTGTATT 155

QY 191 CGTT 194

Db 154 AATT 151

RESULT 61

US-09-783-590-11050

Sequence 11050, Application US/09783590

Patent No. US20020110850A1

GENERAL INFORMATION:

APPLICANT: Dillion, Patrick J.

APPLICANT: Haseltine, William A.

APPLICANT: Li, Haodong

APPLICANT: Rosen, Steven M.

APPLICANT: Ruben, Steven M.

TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2

FILE REFERENCE: PO-16.2C1

CURRENT APPLICATION NUMBER: US/09/783,590

CURRENT FILING DATE: 2000-02-15

PRIOR APPLICATION NUMBER: 08/420,856

PRIOR FILING DATE: 1995-04-12

PRIOR APPLICATION NUMBER: 08/346,731

PRIOR FILING DATE: 1994-11-21

NUMBER OF SEQ ID NOS: 12485

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 11050

LENGTH: 468

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature


```

, LOCATION: (5)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (62)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (133)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (222)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (273)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (283)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (291)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (295)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (313)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (341)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (343)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (409)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (434)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (438)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (439)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (441)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (442)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (443)
, OTHER INFORMATION: n equals a,t,g, or c
, US-09-783-590-11050

```

RESULT 62
US-09-995-898A-29
; Sequence 29, Application US/0995898A
; Publication No. US20030027253A1
; GENERAL INFORMATION:

```

> APPLICANT: Presnell, Scott R.
> APPLICANT: Xu, Wenfeng
> APPLICANT: No. US20030027253Alak, Julia E.
> APPLICANT: Whitmore, Theodore E.
> APPLICANT: Grant, Francis J.
> TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
> FILE REFERENCE: 00-108
> CURRENT APPLICATION NUMBER: US/09/995,898A
> CURRENT FILING DATE: 2001-11-28
> PRIOR APPLICATION NUMBER: US 60/253,561
> PRIOR FILING DATE: 2000-11-28
> PRIOR APPLICATION NUMBER: US 60/267,211
> PRIOR FILING DATE: 2001-02-07
> NUMBER OF SEQ ID NOS: 50
> SOFTWARE: FastSEQ for Windows Version 3.0
> SEQ ID NO 29
> LENGTH: 633
> TYPE: DNA
> ORGANISM: Artificial Sequence
> FEATURE:
> OTHER INFORMATION: Degenerate polynucleotide
> FEATURE:
> NAME/KEY: misc_feature
> LOCATION: (1)...(633)
> OTHER INFORMATION: n = A,T,C or G
> US-09-995-898A-29

```

	Query Match	12.4%;	Score 25.6;	DB 9;	Length 633;
	Best Local Similarity	28.9%;	Pred. No. 27;	Mismatches	37; Conservative 24;
	Gaps	0;	Indels	67;	
Qy	1 AATTGAGCTCGGTACCATACCTGCATATGCTCAAGAACAGTTAAAGTGCTGTGCTA	60			
Dd	264 RAARCARGAYTNTAYAAYAAATTYYAARGGNMGNGTNACNGTNNSCNCNWNSNA	323			
Qy	61 CTAAGCCAGGTCCTTGTCCTATTCTTGTTCCGCTATGCTTAAACCACCTAAC	120			
Dd	324 RWSNCMTGGGTGARWNSGARTAYYNTGYATYYTNTTYGARTNGARCCNGCNC	383			
Qy	121 GTTGTTTTG	128			
Dd	384 NGTNYTNG	391			

```

RESULT 63
US-09-910-943-715/c
; Sequence 715, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/1G148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 715
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(729)
; OTHER INFORMATION: n may be a o r c o r t/u
US-09-910-943-715

```

Query Match	12.4%	Score 25.6;	DB 10;	Length 729;
Best Local Similarity	50.8%;	Pred. No. 28;		
Matches	61;	Conservative	0;	Mismatches 59;
				Indels 0;
Gaps				0;
Qy	75	TGTCCTATTATCTGTGATTCGTTGGCTGATGTTAAACCCACCTAACCGTTGTTTGAAGGAC	134	

Sun Feb 16 09:12:56 2003

Db 279 TCTCGGAGTAGTCGCTTAGTTTCGGGTAGTCAAACTGAAATTCGTTTTCCTTGATTATG 220
QY 135 ACTGATTGTCAGGATCAAAAGTCTGTGAAGGTTCTCGCGGTATGCGCTTGTTCGTT 194
Db 219 ACTGGCTAGACCATATGCAAAAGGTACGAGGTAAAGTCTCTGCTTTTGTGCTTAGTT 160

RESULT 64

US-09-815-242-8813
; Sequence 8813, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8813
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)

US-09-815-242-8813
Query Match 12.4%; Score 25.6; DB 10; Length 942;
Best Local Similarity 59.7%; Pred. No. 32;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 101 TATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTCGCGGTATCAAAAAGTG 160
Db 806 TATGCTATATCGATATATCGATGCTAGATGAATATGTTTGAAGGCGCATATAAAAAA 865

QY 161 CTGTGAAGGTTTC 172
Db 866 CTGCAAGGGTC 877

RESULT 65

US-09-815-242-8969
; Sequence 8969, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8969
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)

US-09-815-242-8969

Query Match 12.4%; Score 25.6; DB 10; Length 942;
Best Local Similarity 59.7%; Pred. No. 32;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 101 TATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTCGCGGTATCAAAAAGTG 160
Db 806 TATGCTATATCGATATATCGATGCTAGATGAATATGTTTGAAGGCGCATATAAAAAA 865

QY 161 CTGTGAAGGTTTC 172

Db 866 CTGCAAGGGTC 877

RESULT 66

US-09-995-898A-3
; Sequence 3, Application US/09995898A
; Publication No. US20030027253A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: No. US20030027253A1ak, Julia E.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Grant, Francis J.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
; FILE REFERENCE: 00-108
; CURRENT APPLICATION NUMBER: US/09/995,898A
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/253,561
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/267,211
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of SEQ ID NO:2
; FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(1473)
; OTHER INFORMATION: n = A,T,C or G
US-09-995-898A-3

Query Match      12.4%; Score 25.6; DB 9; Length 1473;
Best Local Similarity 28.9%; Pred.No. 40;
Matches 37; Conservative 24; Mismatches 67; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGTCTA 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 264 RAARCARGAYTNTAYAAAYAAARTTAAAGGNGMNGTNGMGNACNGTWSNCCNWSNAA 323
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 61 CTAAGCAGGTTCTTCTCCTATTATCTTGTTCGGCTATGTTAAACCCACCTAAC 120
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 324 RWSNCCNTGGTNGARWSNGARTAYTNGAYTAYTNTTYGARGTNGARCCNCCNCC 383
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 121 GTTGTGTTG 128
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 384 NGTNYTNG 391
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 67
US-09-995-898A-28
; Sequence 28, Application US/09995898A
; Publication No. US20030027253A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: No. US20030027253Alak, Julia E.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Grant, Francis J.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
; FILE REFERENCE: 00-108
; CURRENT APPLICATION NUMBER: US/09/995,898A
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/253,561
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/267,211
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate Polynucleotide sequence of SEQ ID
; OTHER INFORMATION: NO:19
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1560)
; OTHER INFORMATION: n = A,T,C or G
US-09-995-898A-28

Query Match      12.4%; Score 25.6; DB 9; Length 1560;
Best Local Similarity 28.9%; Pred.No. 41;
Matches 37; Conservative 24; Mismatches 67; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGTCTA 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 264 RAARCARGAYTNTAYAAAYAAARTTAAAGGNGMNGTNGMGNACNGTWSNCCNWSNAA 323
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 61 CTAAGCAGGTTCTTCTCCTATTATCTTGTTCGGCTATGTTAAACCCACCTAAC 120
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 324 RWSNCCNTGGTNGARWSNGARTAYTNGAYTAYTNTTYGARGTNGARCCNCCNCC 383
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 121 GTTGTGTTG 128
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 384 NGTNYTNG 391
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 68
```

```
US-09-764-864-359
; Sequence 359, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 359
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-359

Query Match      12.4%; Score 25.6; DB 10; Length 2211;
Best Local Similarity 62.5%; Pred.No. 48;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCGGTATGGCTTGT 190
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2002 GAACGTGTGTTGCCAAGGTATTTAAAGGGCTTTCTCGCTCTCTCTTTGATTATT 2061
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 191 CGTT 194
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2062 AATT 2065
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 69
US-09-764-877-3301
; Sequence 3301, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3301
; LENGTH: 6134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3301

Query Match      12.4%; Score 25.6; DB 10; Length 6134;
Best Local Similarity 49.3%; Pred.No. 79;
Matches 67; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 10 TCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGTCTACTAAGCCAG 69
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2631 TCTTTACCTACCTAGCCCAATTGGTTAAATCGGTTCTTCTTCAGAGTGTAGTACAATTT 2690
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 70 GTTCTTGTCCTATTATCTTGTATGTTGGCGGTATGTTAAACCCACCTAACCGTTGTTGA 129
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2691 GTTTTAAAGCTTTTTTTTTTTTTTTTGGTAGCTATAATCATCCCTGTTAAAGCCTTCCTGAG 2750
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 130 AGGACACTGATTGTCC 145
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2751 ATCACCAAAATGTCC 2766
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 70
US-09-070-927A-138
; Sequence 138, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
```


Sun Feb 16 09:12:56 2003

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (450)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (451)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (453)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (454)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (455)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (456)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (462)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (474)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (493)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (494)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-7280

Query Match 12.3%; Score 25.4; DB 10; Length 504;
Best Local Similarity 55.3%; Pred. No. 28;
Matches 47; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 66 CCAGGTTCTTGTCCATTAATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGT 125
DB 294 CCAGGTTCTATGTCCTCTCCACACGTGTATGATGATAGTTTANCCGGTTCCACCTGTG 235
QY 126 TTGAAGGACACTGATTCACAGTA 150
DB 234 CTGGGAACACTGCTCTCCGGGGA 210

RESULT 74
US-09-822-830A-205
Sequence 205, Application US/09822830A
Patent No. US20020142952A1
GENERAL INFORMATION:
APPLICANT: Genetics Institute, Inc.
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6402
CURRENT APPLICATION NUMBER: US/09/822,830A
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,604
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 631
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 205
LENGTH: 1026
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-830A-205

Query Match 12.3%; Score 25.4; DB 10; Length 1026;
Best Local Similarity 52.3%; Pred. No. 39;
Matches 56; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 23 TGCATATGCTCAAGAACCCAGTTAAAGGTCTGTCTACTAAGCCAGGTTCTTCTCTCTAT 82
DB 195 TGCATTTGCTTATCTCTGTGCGAGAACCAATCTCTCTCCACCTATCCCTTCTCTCC 254
QY 83 TATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGA 129
DB 255 CCACCTTGACTGCTCTCTCTCTCAAAATCTTAGCTTCAGATGA 301

RESULT 75
US-08-927-939-55
Sequence 55, Application US/08927939
Patent No. US20010006640A1
GENERAL INFORMATION:
APPLICANT: Grainger, David J.
APPLICANT: Tatalick, Lauen Marie
TITLE OF INVENTION: Compounds and methods to inhibit or
TITLE OF INVENTION: augment an inflammatory response.
FILE REFERENCE: 295, 022US1
CURRENT APPLICATION NUMBER: US/08/927,939
CURRENT FILING DATE: 1997-09-11
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 55
LENGTH: 1041
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (18)...(338)
US-08-927-939-55

Query Match 12.3%; Score 25.4; DB 8; Length 1041;
Best Local Similarity 64.4%; Pred. No. 40;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 60 ACTAAGCCAGGTTCTTGTCTATTTATCTTGTGCTGCTATGTTAAACCCACCTAA 118
DB 562 ACTTACCTGCACACTGCTCTATTTATTTATTTTGAATGTCACCCCAAGTTA 620

Search completed: February 16, 2003, 00:33:05
Job time : 376 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 21:13:02 ; Search time 194 Seconds
(without alignments)
2391.297 Million cell updates/sec

Title: 09-833799-13A

Perfect score: 206

Sequence: 1 AATTCGACTCGGTACCATTA.....GTTTCGTTCCACAATAATAG 206

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N Geneseq 101002.*

```

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	206	100.0	206	11	Sequence encoding
2	127	61.7	8598	19	ppIC9/ELF25L cDNA
3	125.4	60.9	177	15	Elafin derivative
4	125.4	60.9	177	15	Elafin derivative
5	125.2	60.8	480	24	cDNA encoding colo
6	124.6	60.5	434	21	Human head/neck tu
7	124.6	60.5	504	11	Sequence encoding
8	124.6	60.5	571	24	Human colon cancer
9	123.8	60.1	177	15	Elafin derivative

10	123.2	59.8	2309	24	ABK83819	Human cDNA differe
11	123.2	59.8	2309	24	ABL65816	Lung cancer relate
12	123.2	59.8	2309	24	ABL66488	Lung cancer relate
13	121	58.7	321	13	AAQ28757	Partial sequence o
14	66.2	32.1	737	15	AAQ44862	SPAI gene. Sus sc
15	45.4	22.0	321	24	ABK88016	DNA encoding human
16	45.4	22.0	1525	24	ABK88022	DNA sequence encod
17	45.4	22.0	1525	24	ABK88025	DNA sequence encod
18	42.6	20.7	321	21	AAC97529	DNA encoding a pro
19	39	18.9	194	10	AAN90354	Sequence encoding
20	38	18.4	656	24	ABO60721	Human colon cancer
21	34.6	16.8	396	24	ABO73674	Murine SLPI sense
22	34.6	16.8	396	24	ABO73677	Murine SLPI antis
23	34.6	16.8	684	20	AAI18516	Mouse IMC carcinom
24	34.6	16.8	691	20	AAI18514	Mouse IMC carcinom
25	34.6	16.8	1114	20	AAI18515	Mouse IMC carcinom
26	33	16.0	416	22	AAI18195	Human polynucleoti
27	32.2	15.6	324	7	AAN60463	Synthetic sequence
28	32.2	15.6	324	7	AAN60464	Synthetic sequence
29	32.2	15.6	324	20	AAI16194	Serine protease in
30	32.2	15.6	324	20	AAI16236	Serine protease in
31	32.2	15.6	324	21	AAC97527	DNA encoding a ser
32	32.2	15.6	324	21	AAC97528	Secretory leukocyt
33	32.2	15.6	324	21	AAC97579	DNA encoding a ser
34	32.2	15.6	324	22	AAI67488	Recombinant serine
35	32.2	15.6	324	22	AAI67489	Recombinant secret
36	32.2	15.6	420	20	AAI16272	DNA sequence of om
37	32.2	15.6	436	21	AAC97624	DNA encoding ompa-
38	32.2	15.6	436	22	AAI67583	DNA sequence codin
39	32.2	15.6	460	20	AAI16271	DNA sequence of om
40	32.2	15.6	460	21	AAC97622	DNA encoding Ompa
41	32.2	15.6	460	22	AAI67582	DNA sequence codin
42	32.2	15.6	546	23	AAS51617	Staphylococcus aur
43	32.2	15.6	549	23	AAS54717	Staphylococcus aur
44	31.6	15.3	16605	22	AAK65489	Human immune/haema
45	31.6	15.3	16605	22	AAK87049	Human immune/haema
46	30.4	14.8	612	23	ABV53340	Human prostate exp
47	30.4	14.8	3496	22	ABAI18664	Human prostate exp
48	30.4	14.8	160771	24	AQ88179	Human osteoblast d
49	30.4	14.8	580073	18	AAI58840	Mycoplasma genital
50	30.2	14.7	450	21	AAC07910	Human secreted pro
51	30.2	14.7	1489	22	AAH64337	Human secreted pro
52	30.2	14.7	36159	24	ABN85329	Human kinase gene.
53	29.8	14.5	344	22	AAK53717	Murine replication
54	29.4	14.3	5150	18	AAI68952	Blackcurrant fruit
55	29.2	14.2	756	24	ABO89443	Human prostate exp
56	29	14.1	549	22	AAS00193	S. aureus DNA enco
57	29	14.1	549	23	AAS55051	Staphylococcus aur
58	29	14.1	549	23	AAS55319	Staphylococcus aur
59	29	14.1	32145	22	AAK68491	Human immune/haema
60	29	14.1	32145	22	AAK68575	Human immune/haema
61	29	14.1	2365589	24	ABA90521	Genomic sequence o
62	28.8	14.0	1677	21	AAC56088	Eucalyptus grandis
63	28.4	13.8	464	22	ABAI3490	Human breast cell
64	28.4	13.8	464	22	ABAI3490	Human foetal liver
65	28.4	13.8	464	22	ABAI3490	Human foetal liver
66	28.4	13.8	464	22	ABAI3490	Human foetal liver
67	28.4	13.8	464	22	AAK02208	Human brain expres
68	28.4	13.8	464	22	AAK27655	Human bone marrow
69	28.4	13.8	464	22	AAI12235	Probe #2168 for ge
70	28.4	13.8	464	22	AAI133589	Probe #2275 used t
71	28.4	13.8	464	22	AAI02148	Probe #2139 used t
72	28.4	13.8	658	21	AAA16089	Human genome-deriv
73	28.4	13.8	1636	24	ABK34741	Human cDNA for nov
74	28.4	13.8	1656	22	AAV60810	Human secreted pro
75	28.4	13.8	1656	22	AAV60810	Human secreted pro
76	28.4	13.8	1656	22	AAV60810	Human secreted pro
77	28.4	13.8	1656	22	AAV60810	Human secreted pro
78	28.4	13.8	1656	22	AAV60810	Human secreted pro
79	28.4	13.8	1656	22	AAV60810	Human secreted pro
80	28.4	13.8	1656	22	AAV60810	Human secreted pro
81	28.4	13.8	1656	22	AAV60810	Human secreted pro
82	28.4	13.8	1656	22	AAV60810	Human secreted pro

c 83 28.4 13.8 2664 19 AAV59749 Human secreted pro
 c 84 28.4 13.8 125910 21 AAC54370 Human KCNQ5 (KCNQ
 c 85 28.2 13.7 1473 23 ABL20271 Drosophila melanog
 c 86 28.2 13.7 1830 24 AAL45900 A thaliana recepto
 c 87 28.2 13.7 4074 23 ABL20270 Drosophila melanog
 c 88 28.2 13.7 4488 23 ABL02548 Drosophila melanog
 c 89 28 13.6 561 22 AAK91348 Human digestive sy
 c 90 28 13.6 1106 22 AAD12603 Human protein havi
 c 91 28 13.6 1330 22 AAH34853 Human colon cancer
 c 92 28 13.6 1458 22 ABA83379 Human secreted pro
 c 93 28 13.6 2054 24 ABQ91970 Human NF-kB activa
 c 94 28 13.6 2054 24 ABQ91971 Human NF-kB activa
 c 95 28 13.6 2090 22 AAI59408 Human polynucleoti
 c 96 28 13.6 2291 22 AAI61194 Human polynucleoti
 c 97 28 13.6 1830121 17 AA42063 Haemophilus influ
 c 98 27.8 13.5 338 22 AAK59445 Human immune/haema
 c 99 27.8 13.5 626 24 ABQ53920 Oligonucleotide fo
 c 100 27.8 13.5 626 24 ABQ53921 Oligonucleotide fo

ALIGNMENTS

RESULT 1
 ID AAQ06819 standard; DNA; 206 BP.
 AC AAQ06819;
 DT 06-MAR-1991 (first entry)
 DE Sequence encoding human leukocyte elastase inhibiting peptide.
 KW Emphysema; respiratory distress syndrome; atherosclerosis;
 KW arthritis; cystic fibrosis; leukaemia.
 OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 30...203
 FT /*tag= a
 FT /label= Elastase inhibitor

PN EP402068-A.
 XX
 PD 12-DEC-1990.
 XX
 PF 04-JUN-1990; 90EP-0306037.
 XX
 PR 02-NOV-1989; 89GB-0024717.
 PR 09-JUN-1989; 89GB-0013346.
 PR 09-JUN-1989; 89GB-0013349.
 PR 25-SEP-1989; 89GB-0021613.
 XX
 PA (ICIL) IMPERIAL CHEM INDS PLC.
 XX
 PI Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
 XX WPI; 1990-370081/50.
 DR P-PSDB; AAR08217.
 XX

PT Human leukocyte elastase inhibitor is genetically engineered
 PT polypeptide for treatment of inflammatory, pulmonary and skin
 PT conditions
 XX
 PS Disclosure; Fig 13; 45pp; English.

CC Gene product has inhibitory activity against human leukocyte
 CC elastase (HLE) and may be isolated from psoriatic scales. The
 CC sequence may be expressed from a plasmid transformed expression
 CC system and may be useful in the prevention of tissue damage
 CC associated with emphysema, adult respiratory distress syndrome,
 CC psoriasis and bullous dermatoses. Other treatable conditions

CC include atherosclerosis, cystic fibrosis, bronchitis and acute
 CC non-lymphoblastic leukaemia.
 CC Abs raised to the polypeptide may be used in detection.
 XX
 SQ Sequence 206 BP; 51 A; 46 C; 44 G; 65 T; 0 other;
 Query Match 100.0%; Score 206; DB 11; Length 206;
 Best Local Similarity 100.0%; Pred. No. 5.6e-60;
 Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATTCAGCTCGGTACCATACCTGCATATGCTCAAGAACCCAGTTAAAGTCTCTGTGCTA 60
 DB 1 AATTCAGCTCGGTACCATACCTGCATATGCTCAAGAACCCAGTTAAAGTCTCTGTGCTA 60
 QY 61 CTAAGCCAGGTTCTTGTCTCTATTATCTTGATTCTGCTATGTTAAACCCACCTAACC 120
 DB 61 CTAAGCCAGGTTCTTGTCTCTATTATCTTGATTCTGCTATGTTAAACCCACCTAACC 120
 QY 121 GTTGTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGTGTTGAAAGTTCTCTCGGTA 180
 DB 121 GTTGTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGTGTTGAAAGTTCTCTCGGTA 180
 QY 181 TGGCTTGTGTTCTTCCACATAAATAG 206
 DB 181 TGGCTTGTGTTCTTCCACATAAATAG 206

RESULT 2

AAV28852
 ID AAV28852 standard; cDNA to mRNA; 8598 BP.

AC AAV28852;
 XX
 DT 04-AUG-1998 (first entry)
 XX
 DE pPIC9/ELF25L cDNA construct containing the elafin 25L gene.
 XX
 KW Elafin; elafin 25L gene; Pichia pastoris; expression vector;
 KW alcohol oxidase 1; controlling region; promoter; ds.
 XX
 OS Synthetic.
 OS Pichia pastoris.

FH Key Location/Qualifiers
 FT CDS 949..1377
 FT /*tag= a
 FT sig_peptide 949..1203
 FT /*tag= b
 FT mat_peptide 1204..1374
 FT /*tag= c

PN JP10127292-A.
 XX
 PD 19-MAY-1998.
 XX
 PF 31-OCT-1996; 96JP-0304233.
 XX
 PR 31-OCT-1996; 96JP-0304233.
 XX
 PA (TSUR) TSUMURA & CO.

DR WPI; 1998-340667/30.
 DR P-PSDB; AAW57237.

XX New elafin expression vector - used to transform microbial host(s)
 PT for production of elafin commercially
 XX
 PS Example 1; Page 11-15; 21pp; Japanese.

CC The present sequence represents a pPIC9/ELF25L cDNA construct containing
 CC the elafin 25L gene, used in an example of the present invention. The
 CC present invention describes: (1) an elafin (EL) expression vector
 CC comprising a gene encoding EL and a gene encoding alcohol oxidase 1

Db 1 GCACAGGAACCAAGTTAAAGGTCCTCGGTCTCGACCAAAACCGGGCTCTTCCCGCATTCCTG 60
Qy 90 ATTCGTCGCTATGTTAAACCCACCTAAACCGTTGTTGAAGGACACTGATTTGTCAGGT 149
Db 61 ATTCGTCGCTATTTCTGAACCCCGGACCGTTGTTCTGAAGACACTGACTGCCGGGT 120
Qy 150 ATCAAAAAGTCTGTGAAGGTTCTCGGGTAAGCTTGTTCGTTCCACAATAAT 204
Db 121 ATCAAAAAGTCTGTGAAGGTTCTTCGGGTATGCGATGCTTCGTTCCGCGAGTAGT 175

RESULT 5
ABK45842/c
ID ABK45842 standard; cDNA; 480 BP.

XX AC ABK45842;
XX DT 05-JUN-2002 (first entry)
XX DE cDNA encoding colon tumour protein, SEQ ID No 1393.
XX KW Human; colon tumour; vaccine; colon cancer; immunogenic;
XX KW immunotherapy; gene; ss.
XX OS Homo sapiens.
XX PN WO200212328-A2.
XX PD 14-FEB-2002.
XX PF 31-JUL-2001; 2001WO-US24218.
XX PR 03-AUG-2000; 2000US-223283P.
XX PR 28-MAR-2001; 2001US-279763P.
XX PR 29-JUN-2001; 2001US-302051P.
XX PA (CORI-) CORIXA CORP.

XX PI King GE, Meagher MJ, Xu J, Secretist H;
XX DR WPI; 2002-241739/29.
XX PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers
PT for the progression of cancer -

XX PS Claim 1; SEQ ID No 1393; 147pp; English.

XX CC The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer.
CC ABK4450-ABK46237 represent coding sequences of human colon tumour
CC proteins of the invention.

CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
CC for this patent did not form part of the printed specification but was
CC supplied by the European Patent Office.

XX SQ Sequence 480 BP; 113 A; 105 C; 139 G; 120 T; 3 other;

Query Match 60.8%; Score 125.2; DB 24; Length 480;
Best Local Similarity 81.0%; Pred. No. 2e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTCTGTCCTACTAAGCCAGGTTCTTGCTTATATCT 87
Db 371 AAGCGCAAGAGCCAGTCAAGGTCAGTNTCCACTAAGCCTGGCTCTGCCCATATCT 312

Qy 88 TGATTCTGTTGGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGACACTGATTTGCCAG 147
Db 311 TGATCCGGTGGCCATGTTGAATCCCCCTAAACCGCTGCTTGAAGATACTGACTGCCAG 252
Qy 148 GTATCAAAAAGTCTGTGAAGGTTCTTCGCGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 251 GAATCAAGAAGTCTGTGAAGGCTCTTCGGGATGGCCTGTTTCGTTCCCGAGTAGAG 193

RESULT 6
AAC68808/c
ID AAC68808 standard; cDNA; 434 BP.

XX AC AAC68808;
XX DT 21-FEB-2001 (first entry)
XX DE Human head/neck tumour related protein partial coding sequence #5.
XX KW Human head/neck tumour; lung cancer; vaccine; cancer therapy; ss.
XX OS Homo sapiens.
XX PN WO200065053-A2.
XX PD 02-NOV-2000.
XX PF 21-APR-2000; 2000WO-US10687.
XX PR 23-APR-1999; 99US-0130906.
XX PR 20-APR-2000; 2000US-0533870.
XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Dillon DC;
XX DR WPI; 2000-687345/67.

XX PT Novel polypeptides comprising immunogenic portion of head and neck
PT tumour protein useful for treating, diagnosing and monitoring cancer
PT such as head, neck and lung cancer -

XX PS Claim 3; Page 71; 77pp; English.

XX CC The present invention relates to a number of nucleic acid sequences which
CC encode proteins associated with head, neck and lung tumours. These
CC tumours are often not diagnosed until they have spread, and, of those who
CC survive, most must endure alterations in facial and neck appearance as
CC well as changes in speech, sight, smell, chewing, swallowing and taste
CC perception. The coding sequences given, and the proteins they encode, can
CC be used in the diagnosis, treatment and vaccination against cancer,
CC particularly papillary and follicular carcinomas, papillary tumours,
CC follicular adenoma, parathyroid hyperplasia, parotid cancer, lip cancer,
CC squamous cell cancer of the tongue, oral tongue cancers and larynx
CC cancer.

XX SQ Sequence 434 BP; 103 A; 103 C; 133 G; 94 T; 1 other;

Query Match 60.5%; Score 124.6; DB 21; Length 434;
Best Local Similarity 81.0%; Pred. No. 3.1e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAAGTCTGTCCTACTAAGCCAGGTTCTTGCTTATATCT 87
Db 320 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCGGCTCTGCCCATATCT 261

Qy 88 TGATTCTGTTGGCTATGTTAAACCCACCTAAACCGTTGTTTGAAGGACACTGATTTGCCAG 147
Db 260 TGATCCGGTGGCCATGTTGAATCCCCCTAAACCGCTGCTTGAAGATACTGACTGCCAG 201

Qy 148 GTATCAAAAAGTCTGTGAAGGTTCTTCGCGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 200 GAATCAAGAAGTCTGTGAAGGCTCTTCGGGATGGCCTGTTTCGTTCCCGAGTAGAG 142

Db 216 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACACTGACCTGCCAG 275

RESULT 7
AAQ06820

ID AAQ06820 standard; cDNA; 504 BP.

XX AC AAQ06820;

XX DT 06-MAR-1991 (first entry)

XX DE Sequence encoding human leukocyte elastase inhibiting peptide.

XX KW HLE; Emphysema; respiratory distress syndrome; atherosclerosis;
XX KW arthritis; cystic fibrosis; leukaemia; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..157

XX FT /*tag= b

XX FT /label= Upstream in-frame coding sequence

XX CDS 121..328

XX FT /*tag= a

XX FT /label= Elastase inhibitor

XX FT intron 144..180

XX FT /*tag= c

XX PN EP402068-A.

XX PD 12-DEC-1990.

XX PF 04-JUN-1990; 90EP-0306037.

XX PR 02-NOV-1989; 89GB-0024717.

XX PR 09-JUN-1989; 89GB-0013346.

XX PR 09-JUN-1989; 89GB-0013349.

XX PR 25-SEP-1989; 89GB-0021613.

XX PA (ICIL) IMPERIAL CHEM INDS PLC.

XX PI Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
XX WPI; 1990-370081/50.

XX DR P-PSDB; AAR08217.

XX PT Human leukocyte elastase inhibitor - is genetically engineered
XX PT polypeptide for treatment of inflammatory, pulmonary and skin
XX PT conditions

XX PS Disclosure; Fig 16; 45pp; English.

XX CC Gene product has inhibitory activity against human leukocyte
XX CC elastase (HLE) and may be isolated from psoriatic scales. The
XX CC system may be expressed from a plasmid transformed expression
XX CC system and may be useful in the prevention of tissue damage
XX CC associated with emphysema, adult respiratory distress syndrome,
XX CC psoriasis and bullous dermatoses. Other treatable conditions
XX CC include atherosclerosis, cystic fibrosis, bronchitis and acute
XX CC non-lymphoblastic leukaemia.
XX CC Abs raised to the polypeptide may be used in detection.

XX SQ Sequence 504 BP; 110 A; 152 C; 120 G; 122 T; 0 other;

Query Match 60.5%; Score 124.6; DB 11; Length 504;
Best Local Similarity 81.0%; Pred. No. 3.3e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCGAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTATTATCT 87

Db 156 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 215

Qy 88 TGATTCGTCGGCTATGTTAAACCCACTAACCGTTGTTTGAAGGACACTGATTTGCCAG 147

Db 268 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACACTGACCTGCCAG 327

RESULT 8

ABQ58818

ID ABQ58818 standard; cDNA; 571 BP.

XX AC ABQ58818;

XX DT 02-AUG-2002 (first entry)

XX DE Human colon cancer related nucleotide sequence SEQ ID NO:2513.

XX KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX KW genetic analysis; diagnostic; antisense therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200229086-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US30732.

XX PR 02-OCT-2000; 2000US-237271P.

XX PA (FARB) BAYER CORP.

XX PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX PI Thiagalingam A, Lewis ME;

XX DR WPI; 2002-426115/45.

XX PT New isolated nucleic acid that is differentially expressed in cancer
XX PT tissues useful for determining the presence of colon cancer in a cell
XX PT or tissue type, and in antisense therapy -
XX PS Claim 1; Fig 1; 796pp; English.

XX CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
XX CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX CC used in antisense therapy. An antibody immunoreactive with a polypeptide
XX CC encoded by (I) is useful for detecting cancer in a patient sample, and
XX CC for detecting the presence or absence of a polynucleotide encoded by a
XX CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
XX CC from (I) can be used for determining the presence of a nucleic acid which
XX CC hybridises to (I), and for determining the phenotype of cells in a sample
XX CC of cells from a patient. (I) is useful for determining the presence of
XX CC colon cancer in a cell or tissue type, for determining the presence or
XX CC state of other type of cancer, in antisense therapy, to generate
XX CC macroarrays on a solid surface, to identify a chromosome on which the
XX CC corresponding gene resides, and in tissue profiling, forensics, genetic
XX CC analysis, mapping and diagnostic applications. (I) can be used to raise
XX CC antibodies, and to screen for peptide analogues and antagonists.

XX SQ Sequence 571 BP; 136 A; 155 C; 131 G; 145 T; 4 other;

Query Match 60.5%; Score 124.6; DB 24; Length 571;
Best Local Similarity 81.0%; Pred. No. 3.4e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCGAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTGTCTATTATCT 87

Db 208 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCT 267

Qy 88 TGATTCGTCGGCTATGTTAAACCCACTAACCGTTGTTTGAAGGACACTGATTTGCCAG 147

Db 268 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACACTGACCTGCCAG 327

QY 148 GTATCAAAAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206
 |||||
 Db 328 GAATCAAGAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAATAG 386
 |||||
 RESULT 9
 AAQ56675
 ID AAQ56675 standard; DNA; 177 BP.
 XX
 AC AAQ56675;
 XX
 DT 23-SEP-1994 (first entry)
 XX
 DE Elafin derivative Val25 DNA.
 XX
 XX Elafin; derivative; mutation; transformation; E. coli; yeast;
 KW Bacillus subtilis; elastase inhibitor; oxidation; ss.
 KW
 OS Synthetic.
 XX
 PN WO9404697-A.
 XX
 PD 03-MAR-1994.
 XX
 PF 11-AUG-1993; 93WO-JP01133.
 XX
 PR 11-AUG-1992; 92JP-0234085.
 XX
 PA (TSUR) TSUMURA & CO.
 XX
 PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
 XX
 DR WPI; 1994-083211/10.
 DR P-PSDB; AAR48553.
 XX
 PT Recombinant modified elafin with improved oxidation stability -
 PT has pharmaceutical use as an elastase inhibitor
 XX
 PS Claim 1; Page 24; 35pp; Japanese.
 XX
 CC The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
 CC encoding wild type elafin is mutated, inserted into a suitable
 CC vector and then used to transform E. coli, yeast, Bacillus subtilis
 CC or animal cells. The modified elafin is expressed when the
 CC transformed cells are cultured. The modified elafin are drugs with
 CC elastase inhibitor activity. They have improved oxidation stability
 CC over natural elafin and thus retain activity better under oxidative
 CC conditions.
 XX
 SQ Sequence 177 BP; 39 A; 50 C; 48 G; 40 T; 0 other;
 Query Match 60.1%; Score 123.8; DB 15; Length 177;
 Best Local Similarity 81.7%; Pred. No. 4.1e-32;
 Matches 143; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 30 GCTCAAGAACCAAGTTAAAGTCTCTGTCTACTAAGCCAGGTTCTGTTCCTATTATCTTG 89
 |||||
 Db 1 GCACAGGAACCAAGTTAAAGTCCGGTGTGACCAACCGGCTCTTGGCCGATTATCTGT 60
 |||||
 QY 90 ATTCTGGCTATGTTAAACCCACTAACCGTTGTTTGAAGGACACTGATGTCAGGT 149
 |||||
 Db 61 ATCCCGCTGGCTGTCTGTAACCCGCGAACCGTTGTTCTGAAAGACACTGACTGCCCGGT 120
 |||||
 QY 150 ATCAAAAAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAAT 204
 |||||
 Db 121 ATCAAAAAGTGTGTAAGGTTCTCGGTATGGCTTGTTCGTTCCACAATAAT 175
 |||||
 RESULT 10
 ABK83819
 ID ABK83819 standard; cDNA; 2309 BP.
 XX

AC ABR83819;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #390.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US30821.
 XX
 PR 03-OCT-2000; 2000US-237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity
 XX
 PS Claim 1; SEQ ID No 390; 114pp; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GCA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis, renal
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at

```

CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Query Match          59.8%; Score 123.2; DB 24; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.7e-31;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACGAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAAGGTCCTCACTAAGCCGCTGCTCTGCCCATTTATCT 1611

QY 88 TGATTCTGCTGCTATCTTAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147
Db 1612 TGATCCGGTGGCCATGTTGAATCCCTAACCGCTGTTTGAAGATACACTGACTGCCAG 1671

QY 148 GTATCAAAAAGTGTCTGAAGGTTCTGCGGTATGGCTTGTTCGTTCCCAATAA 203
Db 1672 GAATCAAGAAGTGTCTGAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCGCAGTGA 1727

RESULT 11
ABL65816
ID ABL65816 standard; DNA; 2309 BP.
XX
AC ABL65816;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4153.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 26-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.

PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237279P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 4153; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Query Match          59.8%; Score 123.2; DB 24; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.7e-31;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACGAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAAGGTCCTCACTAAGCCGCTGCTCTGCCCATTTATCT 1611

QY 88 TGATTCTGCTGCTATCTTAACCCACCTAACCGTTGTTTGAAGGACACTGATTGCCAG 147
Db 1612 TGATCCGGTGGCCATGTTGAATCCCTAACCGCTGTTTGAAGATACACTGACTGCCAG 1671

QY 148 GTATCAAAAAGTGTCTGAAGGTTCTGCGGTATGGCTTGTTCGTTCCCAATAA 203
Db 1672 GAATCAAGAAGTGTCTGAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCGCAGTGA 1727

RESULT 12
ABL66488
ID ABL66488 standard; DNA; 2309 BP.
XX
XX ABL66488;
XX
AC ABL66488;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4825.

```

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid; CC
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous; CC
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma; CC
 KW gene; ds. CC
 XX Homo sapiens. CC
 OS WO200194629-A2. CC
 XX 13-DEC-2001. CC
 PD 30-MAY-2001; 2001WO-US10838. CC
 XX 05-JUN-2000; 2000US-209473P. CC
 PR 05-JUN-2000; 2000US-209531P. CC
 PR 18-SEP-2000; 2000US-233133P. CC
 PR 18-SEP-2000; 2000US-233617P. CC
 PR 20-SEP-2000; 2000US-234009P. CC
 PR 20-SEP-2000; 2000US-234034P. CC
 PR 20-SEP-2000; 2000US-234052P. CC
 PR 22-SEP-2000; 2000US-234509P. CC
 PR 22-SEP-2000; 2000US-234567P. CC
 PR 25-SEP-2000; 2000US-234923P. CC
 PR 25-SEP-2000; 2000US-234924P. CC
 PR 25-SEP-2000; 2000US-235077P. CC
 PR 25-SEP-2000; 2000US-235082P. CC
 PR 25-SEP-2000; 2000US-235134P. CC
 PR 25-SEP-2000; 2000US-235280P. CC
 PR 26-SEP-2000; 2000US-235637P. CC
 PR 26-SEP-2000; 2000US-235638P. CC
 PR 27-SEP-2000; 2000US-235711P. CC
 PR 27-SEP-2000; 2000US-235720P. CC
 PR 27-SEP-2000; 2000US-235840P. CC
 PR 27-SEP-2000; 2000US-235863P. CC
 PR 28-SEP-2000; 2000US-236028P. CC
 PR 28-SEP-2000; 2000US-236032P. CC
 PR 28-SEP-2000; 2000US-236033P. CC
 PR 28-SEP-2000; 2000US-236034P. CC
 PR 28-SEP-2000; 2000US-236109P. CC
 PR 28-SEP-2000; 2000US-236111P. CC
 PR 29-SEP-2000; 2000US-236842P. CC
 PR 29-SEP-2000; 2000US-236891P. CC
 PR 02-OCT-2000; 2000US-237172P. CC
 PR 02-OCT-2000; 2000US-237173P. CC
 PR 02-OCT-2000; 2000US-237278P. CC
 PR 02-OCT-2000; 2000US-237294P. CC
 PR 02-OCT-2000; 2000US-237295P. CC
 PR 02-OCT-2000; 2000US-237316P. CC
 PR 03-OCT-2000; 2000US-237425P. CC
 PR 03-OCT-2000; 2000US-237598P. CC
 PR 03-OCT-2000; 2000US-237604P. CC
 PR 03-OCT-2000; 2000US-237606P. CC
 PR 03-OCT-2000; 2000US-237608P. CC
 PR 01-NOV-2000; 2000US-244867P. CC
 PR 01-NOV-2000; 2000US-245084P. CC
 XX (AVAL-) AVALON PHARM. CC
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S; CC
 PI Soppet DR, Weaver Z; CC
 DR WPI; 2002-189264/24. CC
 XX Screening for anti-neoplastic agent involves exposing cells to a CC
 PT chemical agent to be tested for anti-neoplastic activity, and CC
 PT determining a change in expression of a gene of a signature gene set - CC
 XX Claim 1; SEQ ID 4825; 44pp; English. CC
 PS The present invention describes a method (M1) for screening for an CC
 CC anti-neoplastic agent. The method involves exposing cells to a chemical CC
 CC agent to be tested for anti-neoplastic activity, determining a change in CC

CC expression of at least one gene (I) of a signature gene set, where (I) CC
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664 CC
 CC to ABL70110), or is at least 95% identical to (S), where a change in CC
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic CC
 CC activity and can be used in gene therapy. M1 can be used for screening CC
 CC an anti-neoplastic agent, and can be used for producing a product which CC
 CC is the data collected with respect to the anti-neoplastic agent as a CC
 CC result of M1, and the data is sufficient to convey the chemical CC
 CC structure and/or properties of the agent. M1 can be used in the CC
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid, CC
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer. CC
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, CC
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine CC
 CC carcinoma, papillary carcinoma and Wilms' tumour. CC
 XX SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other; CC
 Query Match 59.8%; Score 123.2; DB 24; Length 2309; CC
 Best Local Similarity 81.2%; Pred. No. 1.7e-31; CC
 Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0; CC
 QY 28 ATGCTCAAGAACCCAGTTAAAGGTCTGTGTCTACTTAAGCCAGGTTCTTGCTATTATCT 87
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| CC
 1552 AAGCGAAGAGCCAGTCAAAGGTCCAGTCCACTAAGCTGGCTCCTGCCCAATTATCT 1611
 QY 88 TGATTGCTTGCCTATGTTAAACCCACTAACCGTTGTTGAAGGACACTGATTGTCAG 147
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| CC
 1612 TGATCCGTTGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCAG 1671
 QY 148 GTATCAAAAGTGCTGTAAGGTTCTCGGTATGCTTGTTCCTCCACAATAA 203
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| CC
 1672 GAATCAAGAAAGTGCTGTAAGGCTCTTGGGATGCGCTGTTTCGTTCCCAAGTGA 1727
 RESULT 13 CC
 AAQ28757 CC
 ID AAQ28757 standard; DNA; 321 BP. CC
 XX AC AAQ28757; CC
 XX 25-FEB-1993 (first entry) CC
 XX DE Partial sequence of tumour suppressor gene U9. CC
 XX KW Can19; tumour suppressor gene; cancer; therapy; ss. CC
 XX OS Homo sapiens. CC
 XX PN WO9215602-A. CC
 XX PD 17-SEP-1992. CC
 XX PF 28-FEB-1992; 92WO-US01624. CC
 XX PR 28-FEB-1991; 91US-0662216. CC
 XX PA (DAND) DANA FABER CANCER INST INC. CC
 XX PI Sager R; CC
 XX DR WPI; 1992-331663/40. CC
 XX PT Diagnoses and treatment of cancer - using candidate tumor suppressor CC
 XX PT genes or the corresp. antibodies. CC
 XX PS Claim 29; Page 37-38; 54pp; English. CC
 CC An adaptation of the subtractive hybridization technique was used CC
 CC which utilizes a biotinylation-based subtraction procedure instead CC
 CC of hydroxyapatite as previously used. In this procedure, a single CC
 CC strand phagemid cDNA library from normal cell polyA+ mRNA is CC
 CC hybridized with excess biotinylated tumor polyA+ mRNA, and the CC
 CC resulting double stranded sequences are removed by binding to CC

CC streptavidin. The remaining single-stranded phagemid cDNAs are
 CC converted to double-stranded form and used to transform bacterial
 CC host cells. The resulting subtracted cDNA library is differentially
 CC screened with total cDNA from normal and tumor cells. This method
 CC produced some 20 additional cloned cDNAs. Also found by this
 CC method were several genes which, on the basis of the partial DNA
 CC sequences appear to be novel sequences not previously entered
 CC into GENBANK. The portion of the cDNAs so sequenced represents
 CC of the coding region and/or part of the 3' untranslated region
 CC of each cDNA (see Q28749-58).

XX
 SQ Sequence 321 BP; 81 A; 80 C; 84 G; 76 T; 0 other;
 Query Match 58.7%; Score 121; DB 13; Length 321;
 Best Local Similarity 82.2%; Pred. No. 4.6e-31;
 Matches 139; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACAGTAAAGGTCCTGCTACTAAGCCAGGTTCTTGTCTTATCT 87
 DB 142 AAGCGCAAGAGCAGTCAAGAGTCCAGTCTCCACTAAGCGTCTGCTCCCATTTCT 201
 QY 88 TGATTCGTTGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
 DB 202 TGATCCGTCGCGCATGTTGAATCCCTTAACCGCTGCTTGAAGATATGACTGCCCGAG 261
 QY 148 GTATCAAAAAGTCTGTGAAGGTTCTGCGGTATGCGTTGTTTGTTC 196
 DB 262 GAATCAAGAAGTCTGTGAAGGCTCTTGGCGGATGCGCTGTTTCGTTC 310

RESULT 14
 AAQ44862
 ID AAQ44862 standard; cDNA to mRNA; 737 BP.
 XX
 AC AAQ44862;
 DT 17-OCT-1994 (first entry)
 XX
 DE SPAI gene.
 XX
 KW SPAI; sodium ion, potassium ion, ATP-ase inhibitor peptide; cardiac;
 KW vascular disease; PCR; polymerase chain reaction; ss.
 XX
 OS Sus scrofa.

XX
 FH Key Location/Qualifiers
 FT CDS 1..570
 FT /*tag= a
 FT /product= SPAI
 XX
 PN JP06049098-A.
 XX
 PD 22-FEB-1994.
 XX
 PF 29-JUL-1992; 92JP-0202286.
 XX
 PR 29-JUL-1992; 92JP-0202286.
 XX
 PA (EISA) EISAI CO LTD.
 XX
 DR WPI; 1994-097819/12.
 DR P-PSDB; AAR50334.
 XX

PT Sodium ion, potassium ion ATP-ase inhibitor peptide and corresp.
 PT DNA - useful to treat cardiac and vascular disease
 XX
 PS Claim 1; Page 6-7; 8pp; Japanese.

CC The sequence (AAQ44862) encodes the SPAI peptide which is useful
 CC as a drug for treating cardiac and vascular diseases. The peptide
 CC is found as a pre-sequence (AAR50335) which is amplified using
 CC primers (AAQ44863-67)

XX

SQ Sequence 737 BP; 205 A; 192 C; 182 G; 158 T; 0 other;
 Query Match 32.1%; Score 66.2; DB 15; Length 737;
 Best Local Similarity 66.4%; Pred. No. 2.5e-12;
 Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 QY 61 CTAAGCCAGGTTCTTGTCTTATCTTGTTCGTTGGCTATGTTAAACCCACTAAC 120
 DB 428 CTAAGCGTGGCCACTGCCCTAGGATCTTTTCGTTGCCCTGAGCAATCCCTCTACA 487
 QY 121 GTTGTTCGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTCTGAAGTTCTTCGCGTA 180
 DB 488 AGTGTTCGAGAGATTATGACTGTCCAGGGTCAAGAGTGTCTGTGAAGGCTTTTCGCGGA 547
 QY 181 TGGCTTGTTCGTTCCACAATAA 203
 DB 548 AGGATTGTTTGTATCCCAAGTGA 570

RESULT 15
 ABK88016
 ID ABK88016 standard; DNA; 321 BP.
 XX
 AC ABK88016;
 DT 07-OCT-2002 (first entry)
 XX
 DE DNA encoding human secretory leukoprotease inhibitor (SLPI) protein.

XX
 KW Secretory leukoprotease inhibitor; human; gene; ds; protease inhibitor;
 KW malaria; emphysema; asthma; chronic obstructive pulmonary disease; SLPI;
 KW cystic fibrosis; otitis media; otitis externa; HIV; psoriasis; eczema;
 KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
 KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
 KW tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
 KW glomerulonephritis; scleroderma; Alzheimer's disease; hypertension.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT CDS 1..321
 FT /*tag= a
 FT /product= "Secretory leukoprotease inhibitor"
 FT /partial
 FT /note= "No start or stop codon shown"

XX WO20250287-A2.
 XX
 PD 27-JUN-2002.
 XX
 PF 18-DEC-2001; 2001WO-US49256.
 XX
 PR 18-DEC-2000; 2000US-256699P.
 PR 20-NOV-2001; 2001US-331966P.
 XX
 PA (ARRI-) ARRIVA PHARM INC.

XX
 PI Barr PJ, Gibson HL, Pemberton P;
 XX
 DR WPI; 2002-500631/53.
 DR P-PSDB; AAU99874.

XX Novel fusion protein useful for inhibiting protease activity associated
 PT with a disorder such as emphysema, asthma, comprises a first protease
 PT inhibitor comprising alpha 1-antitrypsin and a second protease
 PT inhibitor -

XX Disclosure; Page 32; 134pp; English.

XX This invention relates to a novel fusion protein comprising a first
 CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
 CC active portion and a second protease inhibitor or its functionally
 CC active protein. The fusion proteins of the invention may act as an

CC inhibitor of protease activity. The fusion protein of the invention
 CC is useful for inhibiting protease activity associated with a disorder
 CC such as emphysema, asthma, chronic obstructive pulmonary disease,
 CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
 CC for treating an individual suffering from or at risk for a disease or
 CC disorder involving unwanted protease activity. The proteins are useful
 CC for treating dermatological diseases such as atopic dermatitis, eczema
 CC and psoriasis, in inflammatory responses to viral infection, and for
 CC treating herpes infection, corneal or epidermal ulceration, chronic
 CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
 CC tumour metastasis and tumour angiogenesis, gastric ulceration,
 CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
 CC bacterial infection, Alzheimer's disease, hypertension and muscular
 CC dystrophy. The present sequence represents the DNA encoding the human
 CC secretory leukoprotease inhibitor used to create the fusion protein
 CC of the invention.

XX Sequence 321 BP; 92 A; 67 C; 80 G; 82 T; 0 other;

Query Match 22.0%; Score 45.4; DB 24; Length 321;
 Best Local Similarity 58.5%; Pred. No. 2.1e-05;
 Matches 79; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGCTTATTATCTGATTCGTCGCTATGTTAAACCCACCTAACCGT 122
 DB 178 AAGCCAGGTAAGTGCCAGTACTTACGGTCATGTTTGAATGTTGAACCCACCAACTTC 237
 QY 123 TGTGTAAGGACACTGATTCGACGATATCAAAAAGTCTGTGAAGTTCCTCGGTTATG 182
 DB 238 TGTGAATGACGGTCAATGTAAGAGAGACTTGAAGTGTGTTATGGGTATGTTGTTAAG 297
 QY 183 GCTTGTTCGTTCCA 197
 DB 298 TCCTGTGTTTCCCA 312

RESULT 16
 ABK88022

ID ABK88022 standard; DNA; 1525 BP.

AC ABK88022;

DT 07-OCT-2002 (first entry)

DE DNA sequence encoding SLAP1 fusion protein.

KW Yeast; alpha factor; gene; ds; Alzheimer's disease; SLAP1;
 KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
 KW cystic fibrosis; otitis media; otitis externa; HIV; psoriasis; eczema;
 KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
 KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
 KW tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
 KW glomerulonephritis; scleroderma; hypertension.

OS Homo sapiens.
 OS Synthetic.

Key Location/Qualifiers
 RBS 6..8
 FT /tag= a

FT /standard_name= "Ribosome binding site"

FT 9..1520

FT /tag= b

FT /product= "SLAP1 fusion protein"

FT 12..332

FT /tag= c

FT /note= "SLPI coding region"

FT 333-335

FT /tag= d

FT /note= "linking codon"

FT 336..1517

FT /tag= e

FT /note= "AAT coding region"

XX WO200250287-A2.

XX 27-JUN-2002.

XX 18-DEC-2001; 2001WO-US49256.

XX 18-DEC-2000; 2000US-256699P.

XX 20-NOV-2001; 2001US-331966P.

XX (ARRI-) ARRIVA PHARM INC.

XX Barr PJ, Gibson HL, Pemberton P;

XX WPI: 2002-500631/53.

XX P-PSDB; AU99881.

XX Novel fusion protein useful for inhibiting protease activity associated
 PT with a disorder such as emphysema, asthma, comprises a first protease
 PT inhibitor comprising alpha 1-antitrypsin and a second protease
 PT inhibitor -

XX Example 1; Page 73-73; 134pp; English.

XX This invention relates to a novel fusion protein comprising a first
 CC protease inhibitor comprising an alpha 1-antitrypsin or its functionally
 CC active portion and a second protease inhibitor or its functionally
 CC active protein. The fusion proteins of the invention may act as an
 CC inhibitor of protease activity. The fusion protein of the invention
 CC is useful for inhibiting protease activity associated with a disorder
 CC such as emphysema, asthma, chronic obstructive pulmonary disease,
 CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
 CC for treating an individual suffering from or at risk for a disease or
 CC disorder involving unwanted protease activity. The proteins are useful
 CC for treating dermatological diseases such as atopic dermatitis, eczema
 CC and psoriasis, in inflammatory responses to viral infection, and for
 CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
 CC tumour metastasis and tumour angiogenesis, gastric ulceration,
 CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
 CC bacterial infection, Alzheimer's disease, hypertension and muscular
 CC dystrophy. The present sequence represents the DNA encoding the
 CC SLAP1 fusion protein of the invention.

XX Sequence 1525 BP; 467 A; 286 C; 314 G; 458 T; 0 other;

Query Match 22.0%; Score 45.4; DB 24; Length 1525;

Best Local Similarity 58.5%; Pred. No. 3.7e-05;

Matches 79; Conservative 0; Mismatches 56; Indels 0; Gaps

QY 63 AAGCCAGGTTCTTGCTTATTATCTGATTCGTCGCTATGTTAAACCCACCTAACCGT 122

DB 189 AAGCCAGGTAAGTGCCAGTACTTACGGTCAATGTTGATGTTGAACCCACCAACTTC 248

QY 123 TGTGTAAGGACACTGATTCGACGATATCAAAAAGTCTGTGAAGTTCCTCGGTTATG 182

DB 249 TGTCAATGACGGTCAATGTAAGAGAGACTTGAAGTGTGTTATGGGTATGTTGGTGAAG 308

QY 183 GCTTGTTCGTTCCA 197

DB 309 TCCTGTGTTTCCCA 323

RESULT 17

ABK88025

ID ABK88025 standard; DNA; 1525 BP.

XX AC ABK88025;

XX 07-OCT-2002 (first entry)

XX DNA sequence encoding rSLAP1 fusion protein.

KW rSLAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
 KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
 KW cystic fibrosis; otitis media; otitis externa; HIV; psoriasis; eczema;
 KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
 KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
 KW tumour metastasis; osteoporosis; Paget's disease; scleroderma;
 KW glomerulonephritis; hypertension.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 PH RBS 6..8
 FT /tag= a
 FT /standard_name= "Ribosome binding site"
 FT CDS 9..1520
 FT /tag= b
 FT /product= "rSLAP1 fusion protein"
 FT misc_feature 12..1193
 FT /tag= c
 FT /note= "AAT coding region"
 FT misc_feature 1194..1196
 FT /tag= d
 FT /note= "linking codon"
 FT misc_feature 1197..1517
 FT /tag= e
 FT /note= "SLPI coding region"
 FT
 FT
 XX WO200250287-A2.
 XX
 XX
 XX 27-JUN-2002.
 XX
 XX 18-DEC-2001; 2001WO-US49256.
 XX
 XX 18-DEC-2000; 2000US-256699P.
 XX 20-NOV-2001; 2001US-331966P.
 XX
 XX (ARRI-) ARRIVA PHARM INC.
 XX
 XX Barr PJ, Gibson HL, Pemberton P;
 XX
 XX WPI; 2002-500631/53.
 XX P-PSDB; ANU99884.
 XX
 XX Novel fusion protein useful for inhibiting protease activity associated
 XX with a disorder such as emphysema, asthma, comprises a first protease
 XX inhibitor comprising alpha 1-antitrypsin and a second protease
 XX inhibitor -
 XX
 XX Example 3; Page 89-90; 134pp; English.
 XX
 CC This invention relates to a novel fusion protein comprising a first
 CC protease inhibitor comprising an alpha-antitrypsin or its functionally
 CC active portion and a second protease inhibitor or its functionally
 CC active protein. The fusion proteins of the invention may act as an
 CC inhibitor of protease activity. The fusion protein of the invention
 CC is useful for inhibiting protease activity associated with a disorder
 CC such as emphysema, asthma, chronic obstructive pulmonary disease,
 CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
 CC for treating an individual suffering from or at risk for a disease or
 CC disorder involving unwanted protease activity. The proteins are useful
 CC for treating dermatological diseases such as atopic dermatitis, eczema
 CC and psoriasis, in inflammatory responses to viral infection, and for
 CC treating herpes infection, corneal or epidermal ulceration, chronic
 CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
 CC tumour metastasis and tumour angiogenesis, gastric ulceration,
 CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
 CC bacterial infection, Alzheimer's disease, hypertension and muscular
 CC dystrophy. The present sequence represents the DNA encoding the
 CC rSLAP1 fusion protein of the invention.
 XX
 XX Sequence 1525 BP; 467 A; 287 C; 314 G; 457 T; 0 other;

Query Match 22.0%; Score 45.4; DB 24; Length 1525;
 Best Local Similarity 58.5%; Pred. No. 3.7e-05;
 Matches 79; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
 QY 63 AAGCCAGGTTCTTGCCCTATTATCTTGATTGCTTGGCGCTATGTTAAACCCACCTAACCGT 122
 DB 1374 AAGCCAGGTAAGTGTCAGTTACTTACGGTCAATGTTTGTGTTGAACCCCAACCTTC 1433
 QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTCTGAAGGTTCTCGCGGTATG 182
 DB 1434 TGTGAATGGAGGGTCAATGTAAGAGACATTGAAGTGTGTTGATGTTGTTGTTGAAG 1493
 QY 183 GCTGTTTCGTTCCA 197
 DB 1494 TCTGTGTTTCCCA 1508
 RESULT 18
 AAC97529
 ID AAC97529 standard; DNA; 321 BP.
 XX
 AC AAC97529;
 XX
 DT 27-FEB-2001 (first entry)
 XX
 DE DNA encoding a protease inhibitory peptide.
 XX
 KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.
 XX
 OS Synthetic.
 XX
 PN US6132990-A.
 XX
 PD 17-OCT-2000.
 XX
 PF 07-JUN-1991; 91US-0712354.
 XX
 PR 03-JAN-1989; 89US-0293042.
 PR 06-DEC-1984; 84US-0678822.
 PR 02-DEC-1985; 85US-0803471.
 PR 29-JUL-1986; 86US-0890526.
 PR 30-MAR-1987; 87US-0031846.
 PR 04-AUG-1987; 87US-0082962.
 XX
 PA (AMGE-) AMGEN BOULDER INC.
 XX
 XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX WPI; 2000-678667/66.
 XX
 XX New serine protease inhibitors and DNA sequences for treating a
 XX protease-mediated condition or tissue destruction e.g. emphysema or
 XX tumor invasion and for recombinant production of inhibitors -
 XX
 XX Disclosure; Column 12-13; 47pp; English.
 XX
 CC This invention relates to new purified and isolated mammalian serine
 CC protease inhibitor proteins which comprise at least 8 cysteine residues
 CC and no more than 107 amino acids. The protease inhibitors are capable of
 CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The protease inhibitors have cytostatic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue

CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX Sequence 321 BP; 60 A; 54 C; 60 G; 44 T; 103 other;
 SQ
 Query Match 20.7%; Score 42.6; DB 21; Length 321;
 Best Local Similarity 42.5%; Pred. No. 0.00019;
 Matches 57; Conservative 16; Mismatches 61; Indels 0; Gaps 0;

Qy 63 AAGCAGGTTCTTGTCTATATATCTGATTCGTTGGCTATGTTAAACCCACCTAACCGT 122
 ||:|||||
 Db 178 AARCCNGNAARTGCCNGTNACNTAYGNCARTGYCTNATGCTNAAAYCCNCNAAATY 237
 ||:|||||

Qy 123 TGTGTTGAAGACACTGATGTCAGGATATCAAAAGTGCTGAAGGTTCTCGCGGTATG 182
 ||:|||||
 Db 238 TGNARATGGAYGNCARTGYAARCGNGAYCTNARTGYATGGNATNGGNAAR 297
 ||:|||||

Qy 183 GCTTGTTCGTTCC 196
 ||:|||||
 Db 298 TCNTGYGNTCNC 311
 ||:|||||

RESULT '19

AA090354
 ID AA090354 standard; DNA; 194 BP.

XX AC AA090354;

XX DT 01-NOV-1989 (first entry)

XX Sequence encoding fragment of human polymorphonuclear leukocyte
 XX elastase inhibiting protein.

XX Elastase inhibitor; human polymorphonuclear leukocyte elastase inhibiting
 XX protein.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT misc_feature 1..5

FT /*tag= a

FT misc_feature 7..12

FT /*tag= b

FT misc_feature 25..30

FT /*tag= c

FT misc_feature 56..61

FT /*tag= d

FT misc_feature 120..125

FT /*tag= e

FT misc_feature 185..189

FT /*tag= f

FT misc_feature 190..194

FT /*tag= g

XX WO0906239-A.

XX 13-JUL-1989.

XX 28-DEC-1988; 88WO-JP01342.

XX 28-DEC-1987; 87JP-0330219.

XX (TEIJ) TEIJIN LTD.

XX Sugiyama T, Kamimura T, Masuda K, Okada M, Ohtsuka E;

XX WPI; 1989-220549/30.

XX P-PSDB; AAP90126, AAP90384.

XX Elastase-inhibiting peptide of low serine protease-inhibiting activity
 XX - contains C-terminal portion of human polymorphonuclear leukocyte

PT elastase inhibiting protein.

XX Disclosure; fig 2; 55pp; Japanese.

XX The sequence encodes a fragment of human polymorphonuclear leukocyte
 CC elastase inhibiting protein (Asn 55-Ala 107). See AAP90126 for encoded
 CC peptide. Misc. features a-g are BamHI, SalI, MluI, NdeI, BglII, XhoI and
 CC PstI sites resp. See also AAP90384.

SQ Sequence 194 BP; 48 A; 42 C; 53 G; 51 T; 0 other;

Query Match 18.9%; Score 39; DB 10; Length 194;

Best Local Similarity 55.6%; Pred. No. 0.0026;

Matches 75; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 62 TAAGCCAGGTTCTTGTCTCTATTATCTTGAATTCGTTGGCTATGTTAAACCCACCTAACCG 121
 ||:|||||

Db 33 TAACCCGGTAAATGTCGGTTACATATGTCAGTGTCTGATGCTGAACCGCCGNACTT 92
 ||:|||||

Qy 122 TTGTTGAAGACACTGATTCCTCAGGTATCAAAAAGTGTGTGAAGTTCCTCGGTAT 181
 ||:|||||

Db 93 CTGTGAATGGACGTCAGTCTAAACGAGATCTGAAATGTTGATGGGTATGTGTGTAA 152
 ||:|||||

Qy 182 GGCTTGTTCGTTCC 196
 ||:|||||

Db 153 ATCTTGTGTTCTCC 167
 ||:|||||

RESULT 20

ABQ60721/c

ID ABQ60721 standard; cDNA; 656 BP.

XX AC ABQ60721;

XX DT 02-AUG-2002 (first entry)

XX Human colon cancer related nucleotide sequence SEQ ID NO:4416.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
 XX genetic analysis; diagnostic; antisense therapy; gene; ss.

XX OS Homo sapiens.

XX WO200229086-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US30732.

XX PR 02-OCT-2000; 2000US-237271P.

XX PA (FARB) BAYER CORP.

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;

XX Thiagalingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer
 XX tissues useful for determining the presence of colon cancer in a cell
 XX or tissue type, and in antisense therapy -

XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABQ7893 to ABQ7904 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridises to (I), and for determining the phenotype of cells in a sample

CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate
 CC microarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists.

XX Sequence 656 BP; 96 A; 91 C; 107 G; 256 T; 106 other;

Query Match 18.4%; Score 38; DB 24; Length 656;

Best Local Similarity 65.7%; Pred. No. 0.0088;

Matches 44; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 139 ATTCTCCAGGTATCAAAAGTGTGGAAGTCTCTCGGTATGGCTGTGTTCTGCCAC 198

DB 285 ACTGNNCAGGAATCAAGAANTNTNGAAGGNTCTAGNGGGATGGCCNGATTTGTTCCCC 226

QY 199 AATAATA 205

DB 225 ANNGANA 219

RESULT 21

ABQ73674

ID ABQ73674 standard; DNA; 396 BP.

AC ABQ73674;

DT 04-OCT-2002 (first entry)

DE Murine SLPI sense riboprobe SEQ ID NO:20.

XX Murine; serum amyloid; SAA3; riboprobe; LST-1; follistatin-like 1;
 KW FST1; SLPI; secretory leukocyte protease inhibitor; legumain; SOCS3;
 KW rheumatoid arthritis; CISH3; RAGE; AGER; LY117; hydroxymethyl glutarate;
 KW HMG-1; glucocorticoid leucine zipper; GILZ; PTPN-18; GADD-45A; GADD-45B;
 KW PRSC1; lipocalin 2; Lcn2; glucose phosphate isomerase; GPI; SpIL;
 KW serine protease inhibitor; TSG-6; ss.

XX Mus sp.

XX WO200248310-A2.

XX 20-JUN-2002.

XX 17-DEC-2001; 2001WO-US48968.

XX 15-DEC-2000; 2000US-255861P.

XX (GENE-) GENETICS INST LLC.

PI Pittman DD, Feldman JL, Shields KM, Trepicchio WL;

XX WPI; 2002-583494/62.

XX Determining difference between expression levels of genes
 PT characteristic of rheumatoid arthritis in cell and reference level, by
 PT comparing expression levels of the genes determined in a cell with
 PT reference level -

XX Example 3; Page 88; 296pp; English.

XX The present invention describes a method (M1) for determining the
 CC difference between levels of expression of a number of genes
 CC characteristic of rheumatoid arthritis (RA) in cell and reference levels
 CC of expression of the genes, comprising: (a) providing RNA from a cell;
 CC (b) determining levels of RNA of a number of genes (G) characteristic
 CC of RA including a number of genes selected from SOCS3 (CISH3), RAGE
 CC (AGER), LST-1 (LY117), serum amyloid (SAA) 1-3, hydroxymethyl glutarate
 CC (HMG)-1, S100 A8, A9, and A12, secretory leukocyte protease inhibitor
 CC (SLPI), glucocorticoid leucine zipper (GILZ), PTPN-18, GADD-45A and B,
 CC Legumain (PRSC1), follistatin-like 1 (FST1), lipocalin 2 (Lcn2), glucose

CC phosphate isomerase (GPI), serine protease inhibitor (SpIL), and TSG-6
 CC to obtain the levels of expression of the genes in the cell; and
 CC (c) comparing the levels of expression of the genes in the cell to a
 CC set of reference levels of expression of the genes, to determine the
 CC difference between levels of expression of the number of genes
 CC characteristic of RA in the cell and reference levels of expression of
 CC the genes. M1 is useful for determining whether a subject has or is
 CC likely to develop RA, or for determining whether a therapy for RA is
 CC effective in a subject having RA who is receiving the therapy. M1 is
 CC also useful for determining the stage of RA in a subject, for
 CC determining the efficacy of a therapy in a subject having RA, and for
 CC determining the likelihood of success of a particular therapy in subject
 CC having RA. The present sequence represents a murine SLPI sense riboprobe,
 CC which is used in an example from the present invention.

XX Sequence 396 BP; 94 A; 98 C; 118 G; 86 T; 0 other;

Query Match 15.8%; Score 34.6; DB 24; Length 396;

Best Local Similarity 53.3%; Pred. No. 0.1;

Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0

QY 51 CCTGTGTCTACTAAGCCAGGTTCTTGCTCTATTATCTTGATTCGTGCGCTATGTTAAAC 110

DB 244 CCAGTGTGGAGGAGCCCTGGGAGGTGCGTCAAACTCAGGCAAGATGATGATGCTTAAC 303

QY 111 CCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGT 170

DB 304 CCTCCCAATGTCGCCAGAGGACGGCGAGGTGTGACGGCAATACAGTGTCTGTGAGGGT 363

QY 171 TCTTCGGGTATGGCTTG 187

DB 364 ATATGTGGGAAAGTCTG 380

RESULT 22

ABQ73677/C

ID ABQ73677 standard; DNA; 396 BP.

AC ABQ73677;

DT 04-OCT-2002 (first entry)

DE Murine SLPI antisense riboprobe SEQ ID NO:23.

XX Murine; serum amyloid; SAA3; riboprobe; LST-1; follistatin-like 1;
 KW FST1; SLPI; secretory leukocyte protease inhibitor; legumain; SOCS3;
 KW rheumatoid arthritis; CISH3; RAGE; AGER; LY117; hydroxymethyl glutarate;
 KW HMG-1; glucocorticoid leucine zipper; GILZ; PTPN-18; GADD-45A; GADD-45B;
 KW PRSC1; lipocalin 2; Lcn2; glucose phosphate isomerase; GPI; SpIL;
 KW serine protease inhibitor; TSG-6; ss.

OS Mus sp.

XX WO200248310-A2.

XX 20-JUN-2002.

XX 17-DEC-2001; 2001WO-US48968.

XX 15-DEC-2000; 2000US-255861P.

XX (GENE-) GENETICS INST LLC.

PI Pittman DD, Feldman JL, Shields KM, Trepicchio WL;

XX WPI; 2002-583494/62.

XX Determining difference between expression levels of genes
 PT characteristic of rheumatoid arthritis in cell and reference level, by
 PT comparing expression levels of the genes determined in a cell with
 PT reference level -

XX Example 3; Page 88; 296pp; English.

PI Arakawa H, Morita M, Ohta M;
DR WPI; 1999-080732/07.
XX P-PSDB; AAW98908.
XX
PT Protein associated with cancer metastasis and gene encoding it -
XX useful for screening for potential inhibitors of cancer metastasis
XX
PS Claim 2; Page 41-42; 74pp; Japanese.
XX
CC The present invention provides gene sequences associated with cancer
CC metastasis which are isolated from mouse IMC carcinoma cells by
CC detection of their higher expression in IMC-HM cell lines than in
CC IMC-LM cell lines using differential display of the mRNA in these cells.
CC The gene sequences can be used for the screening of potential inhibitors
CC of cancer metastasis by either: bringing into contact with the cancer
CC metastasis associated protein (CMAP) and determining the degree of
CC binding; or creating a transformant cell line which expresses CMAP and
CC measuring the degree of expression of CMAP using an antibody recognising
CC the protein, either in the presence or absence of the potential
CC inhibitor. IMC-HM cells transformed with antisense CMAP DNA show a
CC lowered ability to metastasise. The present sequence represents a
CC specifically claimed gene sequence from the present invention.
XX
SQ Sequence 691 BP; 160 A; 186 C; 183 G; 162 T; 0 other;
Query Match 16.8%; Score 34.6; DB 20; Length 691;
Best Local Similarity 53.3%; Pred. No. 0.13;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 51 CCTGTGCTCTACTAAGCCAGGTTCTTGTCTTATTTATCTTGTGCGCTATGTTAAAC 110
DB 280 CCAGTGTGGAGGAAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGTATGCTTTAAC 339
QY 111 CCACCTAACCGTTGTTGAAGGACACTGATTGTCAGGTTATCAAAAGTCTGTGAAGGT 170
DB 340 CCTCCCAATGCTGCCAGAGGGACGGCAGTGTACCGCAATATCAAGTGTGTGAGGTT 399
QY 171 TCCTGCGGTATGCTTG 187
DB 400 ATATGTGGGAAGTCTG 416
RESULT 25
AAI81895
ID AAI81895 standard; cDNA; 1114 BP.
AC AAI81895;
XX
DT 05-MAY-1999 (first entry)
XX
DE Mouse IMC carcinoma cell IMC-HA1 clone #8.323 cDNA.
XX
KW Mouse; carcinoma cell; IMC-HA1; cancer; metastasis; CMAP; inhibitor;
KW cancer metastasis associated protein; ss.
XX
OS Mus musculus.
XX
PN WO9845431-A1.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-JP01592.
XX
PR 08-APR-1997; 97JP-0105333.
XX
XX (BANY) BANYU PHARM CO LTD.
XX
PA Arakawa H, Morita M, Ohta M;
PI WPI; 1999-080732/07.
DR
XX Protein associated with cancer metastasis and gene encoding it -
PT useful for screening for potential inhibitors of cancer metastasis
XX
PS Claim 2; Page 44-45; 74pp; Japanese.
XX
CC The present invention provides gene sequences associated with cancer
CC metastasis which are isolated from mouse IMC carcinoma cells by
CC detection of their higher expression in IMC-HM cell lines than in
CC IMC-LM cell lines using differential display of the mRNA in these cells.
CC The gene sequences can be used for the screening of potential inhibitors
CC of cancer metastasis by either: bringing into contact with the cancer
CC metastasis associated protein (CMAP) and determining the degree of
CC binding; or creating a transformant cell line which expresses CMAP and
CC measuring the degree of expression of CMAP using an antibody recognising
CC the protein, either in the presence or absence of the potential
CC inhibitor. IMC-HM cells transformed with antisense CMAP DNA show a
CC lowered ability to metastasise. The present sequence represents a
CC specifically claimed gene sequence from the present invention.
XX
SQ Sequence 1114 BP; 271 A; 299 C; 290 G; 254 T; 0 other;
Query Match 16.8%; Score 34.6; DB 20; Length 1114;
Best Local Similarity 53.3%; Pred. No. 0.15;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 51 CCTGTGCTCTACTAAGCCAGGTTCTTGTCTTATTTATCTTGTGCGCTATGTTAAAC 110
DB 697 CCAGTGTGGAGGAAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGTATGCTTTAAC 756
QY 111 CCACCTAACCGTTGTTGAAGGACACTGATTGTCAGGTTATCAAAAGTCTGTGAAGGT 170
DB 757 CCTCCCAATGCTGCCAGAGGGACGGCAGTGTACCGCAATATCAAGTGTGTGAGGTT 816
QY 171 TCCTGCGGTATGCTTG 187
DB 817 ATATGTGGGAAGTCTG 833
RESULT 26
AAI81895
ID AAI81895 standard; cDNA; 416 BP.
AC AAI81895;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1955.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
DR WPI; 2001-514838/56.
XX P-PSDB; AAO01964.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 1; SEQ ID NO 1955; 1399pp + Sequence Listing; English.
 PS
 XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA00010-AA03910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 416 BP; 102 A; 77 C; 73 G; 163 T; 1 other;
 Query Match 16.0%; Score 33; DB 22; Length 416;
 Best Local Similarity 54.5%; Pred. No. 0.37;
 Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 26 ATATGCTCAAGAACAGTTAAAGGTCCTGCTCTACTAAGCCAGGTCCTGTCCTATTAT 85
 Db 140 ACATGCTGGACCAAACTAGAGAGTCATTAACTTTTAAAAAGAACTGCCACACTGTTT 199
 QY 86 CTGATGCTGGCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTCGCC 145
 Db 200 CTTAATTGGTGTACCACTTTTACATTCACCAGCAATGTGTGAAGATTCCTGTTCTCC 259
 QY 146 A 146
 Db 260 A 260
 RESULT 27
 AAN60463
 ID AAN60463 standard; DNA; 324 BP.
 XX
 AC AAN60463;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE Synthetic sequence capable of directing microbial synthesis of a
 DE serine protease inhibitor having similar properties to protein
 DE isolated from parotid secretions.
 XX
 KW Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
 KW trypsin-; inhibitor; ss.
 XX
 OS Synthetic.
 XX
 FN WO8603519-A.
 XX
 PD 19-JUN-1986.
 XX
 PF 04-DEC-1985; 85WO-US02385.
 XX
 PR 04-DEC-1985; 85WO-US02385.
 PR 06-DEC-1984; 84US-0678822.
 PR 02-DEC-1985; 85US-0803471.
 XX
 PA (SYNE-) SYNERGEN BIOLOG INC.
 XX
 PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX
 WPI; 1986-169458/26.
 DR P-PSDB; AAP60562.
 XX
 XX New synthetic DNA sequences for directing microbial synthesis -
 PT for prodn. of single poly:peptide chain serine protease inhibitor

PT having leukocyte elastase and trypsin inhibitory sites
 XX
 PS Disclosure; Page 14; 59pp; English.
 XX
 CC The sequence directs synthesis of a single chain polypeptide serine
 CC protease-inhibitor, which believed to have at least 2 active sites,
 CC 1 exhibiting leukocyte-elastase-inhibiting properties and the other
 CC exhibiting activity against trypsin. See also AAN60464-69 and AAP60562-
 CC 66.
 XX
 SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;
 Query Match 15.6%; Score 32.2; DB 7; Length 324;
 Best Local Similarity 51.8%; Pred. No. 0.63;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 63 AAGCCAGGTTCTTGTCTATTATCTTGTTCGTCGCTATGTTAAACCCACCTAACCGT 122
 Db 178 AAACCGGTAATGCCCGGTAACTATGGCCAGTGTCTGATGCTGAACCCCGCACTTC 237
 QY 123 TGTTTTAAAGGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGGTTCTTCGGGTATG 182
 Db 238 TCGAAATGAGCGGCCAGTGTAAACGAGATCTGAAATCTGTATGGGTATGTGGGCAAA 297
 QY 183 GCTTGTTCGTTCCACAATAA 203
 Db 298 TCTTGTGTTTCCCGGTAAAA 318
 RESULT 28
 AAN60464
 ID AAN60464 standard; DNA; 324 BP.
 XX
 AC AAN60464;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE Synthetic sequence capable of directing microbial synthesis of a
 DE secretory leukocyte protease-inhibitor.
 XX
 KW Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
 KW trypsin-; inhibitor; ss.
 XX
 OS Synthetic.
 XX
 FN WO8603519-A.
 XX
 PD 19-JUN-1986.
 XX
 PF 04-DEC-1985; 85WO-US02385.
 XX
 PR 04-DEC-1985; 85WO-US02385.
 PR 06-DEC-1984; 84US-0678822.
 PR 02-DEC-1985; 85US-0803471.
 XX
 PA (SYNE-) SYNERGEN BIOLOG INC.
 XX
 PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX
 WPI; 1986-169458/26.
 DR P-PSDB; AAP60563.
 XX
 XX New synthetic DNA sequences for directing microbial synthesis -
 PT for prodn. of single poly:peptide chain serine protease inhibitor
 PT having leukocyte elastase and trypsin inhibitory sites
 XX
 PS Disclosure; Page 15; 59pp; English.
 XX
 CC The sequence directs synthesis of a secretory leukocyte protease-
 CC inhibitor. See also AAN60463, AAN60465-69 and AAP60562. AAP60564-66.
 XX
 SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match	15.8%;	Score 32.2;	DB 7;	Length 324;
Best Local Similarity	51.8%;	Pred. No. 0.63;		
Matches	73;	Conservative	0;	Mismatches 68; Indels 0; Gaps

QY	63	AAGCCAGGTTCTTGCTCTATTATCTTGATTGCTTGCTATGTTAAACCCACCTAACCGT	122
DB	178	AAACCGGTAATGCCCGGTAACTATGCGCAGTGCTGATGCTGAACCCGCGGAACCTTC	237
QY	123	TGTTTGAAGGACACTGATTGTCAGGATATCAAAAGTGCTGTGAAGGTTCTTCGCGGTATG	182
DB	238	TGCGAAATGACGCGCCAGTGTAAACAGAGATCTGAAATGCTGTATGGTATGTGCGCAAA	297
QY	183	GCITGTTTCGTTCCACAATAA	203
DB	298	TCTTGTGTTTCCCGGTAAAA	318

RESULT 29	
AAAX16194	
ID	AAAX16194 standard; DNA; 324 BP.
XX	XX
XX	AC
XX	AAAX16194;
XX	XX
DT	19-APR-1999 (first entry)
DE	Serine protease inhibitor direct manufacturing DNA sequence.
XX	XX
KW	Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW	muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW	acute leukemia; ss.
XX	XX
OS	Homo sapiens.
OS	Synthetic.
XX	XX
PN	US5871956-A.
XX	XX
PD	16-FEB-1999.
XX	XX
PF	22-JUL-1994; 94US-0279056.
XX	XX
PR	30-MAR-1987; 87US-0031846.
PR	03-DEC-1984; 84US-0678822.
PR	29-JUL-1986; 86US-0890526.
PR	03-SEP-1986; 86US-0903471.
PR	06-AUG-1990; 90US-0563832.
PR	22-JUL-1994; 94US-0279056.
XX	XX
PA	(AMGE-) AMGEN INC.
XX	XX
PI	Bandypadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX	XX
WI	WI; 1999-166640/14.
DR	
XX	XX

PT	New DNA sequence encoding mammalian serine protease inhibitor -
PT	useful for recombinantly producing inhibitors with different
PT	specificities and treating diseases such as emphysema, arthritis,
PT	muscular dystrophy, and tumour invasion
XX	XX
PS	Example 1; Column 23; 37pp; English.
XX	XX

CC	The present invention describes a DNA sequence (A) encoding an analog
CC	of a mammalian serine protease inhibitor (B). The DNA sequences and
CC	recombinant methods allow manufacture of a class of inhibitors of e.g.
CC	cathepsin G, elastase, and trypsin, with different specificities. The
CC	recombinant serine protease product can be directed to act
CC	intracellularly or extracellularly and is useful in treating conditions
CC	caused by a disturbance in the native protease/protease inhibitor
CC	balance such as emphysema, arthritis, glomerulonephritis, peridontitis
CC	muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
CC	allows (B) to be recombinantly produced in sufficient quantities and
CC	purities so as to provide economical pharmaceutical compositions. (B)
CC	resistant to heat, acid, and a variety of proteolytic enzymes, is
CC	thermodynamically stable in extracellular conditions, and exhibits a

CC caused by a disturbance in the native protease/protease inhibitor
 CC balance such as emphysema, arthritis, glomerulonephritis, periodontitis,
 CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
 CC allows (B) to be recombinantly produced in sufficient quantities and
 CC purities so as to provide economical pharmaceutical compositions. (B) is
 CC resistant to heat, acid, and a variety of proteolytic enzymes, is
 CC thermodynamically stable in extracellular conditions, and exhibits a
 CC high degree of self assembly forming an active tertiary structure in the
 CC absence of biochemical stimuli. The present sequence represents a DNA
 CC sequence which is used to direct manufacture of recombinant serine
 CC protease inhibitors.

XX SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match 15.6%; Score 32.2; DB 20; Length 324;
 Best Local Similarity 51.8%; Pred. No. 0.63;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 63 AAGCCAGGTTCTTGTCTTATTATCTTGATTCGTTGCGCTATGTTTAAACCCACCTAACCGT 122
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 178 AAACCGGGTAAATGCCCGTAAACCTATGCGCCAGTGTCTGATGCTGAACCCGCGAACTTC 237

Qy 123 TGTTTGAAGGACACTGATTCCTCAGGTATCAAAAGTCTGTGAAGTTCCTCGGTATG 182
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 238 TCGGAAATGGACGCGCAGTGAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAA 297

Qy 183 GCTTGTTCGTTCCACAATAA 203
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 298 TCTTGTGTTTCCCGGTAAAA 318

RESULT 31
 AAC97527

ID AAC97527 standard; DNA; 324 BP.

XX AC AAC97527;

XX DT 27-FEB-2001 (first entry)

XX DE DNA encoding a serine protease inhibitor peptide sequence.

XX KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.

XX OS Synthetic.

XX PN US6132990-A.

XX XX PD 17-OCT-2000.

XX XX PF 07-JUN-1991; 91US-0712354.

XX XX PR 03-JAN-1989; 89US-0293042.

XX XX PR 06-DEC-1984; 84US-0678822.

XX XX PR 02-DEC-1985; 85US-0803471.

XX XX PR 29-JUL-1986; 86US-0890526.

XX XX PR 30-MAR-1987; 87US-0031846.

XX XX PR 04-AUG-1987; 87US-0082962.

XX XX PA (AMGE-) AMGEN BOULDER INC.

XX XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX XX WPI; 2000-678667/66.

XX XX PT New serine protease inhibitors and DNA sequences for treating a
 PT protease-mediated condition or tissue destruction e.g. emphysema or
 XX tumor invasion and for recombinant production of inhibitors -

XX PS Claim 34; Column 64-66; 47pp; English.

XX

CC This invention relates to new purified and isolated mammalian serine
 CC protease inhibitor proteins which comprise at least 8 cysteine residues
 CC and no more than 107 amino acids. The protease inhibitors are capable of
 CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The protease inhibitors have cytostatic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;

Query Match 15.6%; Score 32.2; DB 21; Length 324;
 Best Local Similarity 51.8%; Pred. No. 0.63;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 63 AAGCCAGGTTCTTGTCTTATTATCTTGATTCGTTGCGCTATGTTTAAACCCACCTAACCGT 122
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 178 AAACCGGGTAAATGCCCGTAAACCTATGCGCCAGTGTCTGATGCTGAACCCGCGAACTTC 237

Qy 123 TGTTTGAAGGACACTGATTCCTCAGGTATCAAAAGTCTGTGAAGTTCCTCGGTATG 182
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 238 TCGGAAATGGACGCGCAGTGAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAA 297

Qy 183 GCTTGTTCGTTCCACAATAA 203
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 298 TCTTGTGTTTCCCGGTAAAA 318

RESULT 32

AAC97528

ID AAC97528 standard; DNA; 324 BP.

XX AC AAC97528;

XX XX DT 27-FEB-2001 (first entry)

XX DE Secretory leukocyte protease inhibitor (SLPI) DNA sequence.

XX KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.

XX OS Synthetic.

XX XX PN US6132990-A.

XX XX PD 17-OCT-2000.

XX XX PF 07-JUN-1991; 91US-0712354.

XX XX PR 03-JAN-1989; 89US-0293042.

XX XX PR 06-DEC-1984; 84US-0678822.

XX XX PR 02-DEC-1985; 85US-0803471.

XX XX PR 29-JUL-1986; 86US-0890526.

XX XX PR 30-MAR-1987; 87US-0031846.

XX XX PR 04-AUG-1987; 87US-0082962.

XX XX PA (AMGE-) AMGEN BOULDER INC.

XX XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX XX WPI; 2000-678667/66.

XX New serine protease inhibitors and DNA sequences for treating a
 PT protease-mediated condition or tissue destruction e.g. emphysema or
 PT tumor invasion and for recombinant production of inhibitors -
 XX
 XX Disclosure; Column 5; 47pp; English.
 XX
 CC This invention relates to new purified and isolated mammalian serine
 CC protease inhibitor proteins which comprise at least 8 cysteine residues
 CC and no more than 107 amino acids. The protease inhibitors are capable of
 CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitors of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The protease inhibitors have cytostatic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match 15.6%; Score 32.2; DB 21; Length 324;
 Best Local Similarity 51.8%; Pred. No. 0.63;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 63 AACCCAGGTTCTTGTCTATTATCTTGTGTTGCGCTATGTTAAACCCACCTAACCGT 122
 DB 178 AACCCGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCGCGAACTTC 237
 QY 123 TGTTTGAAGGACACTGATGTCAGGTATCAAAAAGTGTGTAAGGTTCTCTGCGGTATG 182
 DB 238 TGCGAATGACGCGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGCAAA 297
 QY 183 GCTTGTTCGTTCCACAATAA 203
 DB 298 TCTTGTGTTTCCCGGTAAAA 318

RESULT 33
 AAC97579

ID AAC97579 standard; DNA; 324 BP.

XX AAC97579;

XX 27-FEB-2001 (first entry)

XX DNA encoding a serine protease inhibitory peptide.

XX Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.

XX Synthetic.

XX US6132990-A.

XX 17-OCT-2000.

XX 07-JUN-1991; 91US-0712354.

XX 03-JAN-1989; 89US-0293042.

XX 06-DEC-1984; 84US-0678822.

XX 02-DEC-1985; 85US-0803471.

XX 29-JUL-1986; 86US-0890526.

XX 30-MAR-1987; 87US-0031846.

PR 04-AUG-1987; 87US-0082962.

XX (AMGE-) AMGEN BOULDER INC.

XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX WPI; 2000-678667/66.

XX New serine protease inhibitors and DNA sequences for treating a
 PT protease-mediated condition or tissue destruction e.g. emphysema or
 PT tumor invasion and for recombinant production of inhibitors -
 XX

PS Example 2; Column 30-32; 47pp; English.

XX This invention relates to new purified and isolated mammalian serine
 CC protease inhibitor proteins which comprise at least 8 cysteine residues
 CC and no more than 107 amino acids. The protease inhibitors are capable of
 CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitors of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The serine protease inhibitors have cytostatic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.

XX Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match 15.6%; Score 32.2; DB 21; Length 324;
 Best Local Similarity 51.8%; Pred. No. 0.63;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AACCCAGGTTCTTGTCTATTATCTTGTGTTGCGCTATGTTAAACCCACCTAACCGT 122
 DB 178 AACCCGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCGCGAACTTC 237
 QY 123 TGTTTGAAGGACACTGATGTCAGGTATCAAAAAGTGTGTAAGGTTCTCTGCGGTATG 182
 DB 238 TGCGAATGACGCGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGCAAA 297
 QY 183 GCTTGTTCGTTCCACAATAA 203
 DB 298 TCTTGTGTTTCCCGGTAAAA 318

RESULT 34

AAI67488

ID AAI67488 standard; DNA; 324 BP.

XX AAI67488;

XX 11-FEB-2002 (first entry)

XX Recombinant serine protease inhibitor DNA.

XX Serine protease inhibitor protein; recombinant; leukocyte elastase;
 KW trypsin; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..324

XX /tag= a

XX /product= "serine protease inhibitor"

XX US6291662-B1.

```

XX PD 18-SEP-2001.
XX PF 22-SEP-1998; 98US-0158085.
XX PR 30-MAR-1987; 87US-0031846.
XX PR 06-AUG-1990; 90US-0563832.
XX PR 22-JUL-1994; 94US-0279056.
XX PR 05-DEC-1984; 84US-0678222.
XX PR 02-DEC-1985; 85US-0803471.
XX PR 29-JUL-1986; 86US-0890526.
XX PA (AMGE-) AMGEN INC.
XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX DR WPI; 2001-637974/73.
XX DR P-PSDB; AAG65993.
XX PT New DNA sequences, useful in recombinant DNA techniques for directing
XX PT the production of a serine protease inhibitor protein, e.g. leukocyte
XX PT elastase or trypsin -
XX PS Example 1; Column 4; 37pp; English.
XX CC The invention relates to a DNA sequence encoding an analog of a mammalian
XX CC serine protease inhibitor protein. The analog comprises at least eight
XX CC cysteine residues and possesses serine protease inhibitor activity, and
XX CC at least one operational element consisting of a promoter, an operator,
XX CC or a terminator codon. The DNA is useful in recombinant DNA techniques
XX CC for directing the production of a serine protease inhibitor protein,
XX CC e.g. leukocyte elastase or trypsin. The present sequence represents the
XX CC coding strand of a DNA that directs the manufacture of a recombinant
XX CC serine protease inhibitor.
XX SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;

Query Match 15.6%; Score 32.2; DB 22; Length 324;
Best Local Similarity 51.8%; Pred. No. 0.63;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCAGGTTCTTGCTTATTTATCTTGGTTCGTCGCTATGTTAAACCCACCTAACCGT 122
Db 178 AAACCGGGTAAATGCCCGTAACCTATGCCAGTGTCTGATGCTGACCCCGCACTTC 237
QY 123 TGTTTGAAGGACACTGATGTCAGGTATCAAAAGTCTGTAAGGTTCTCGCGGTATG 182
Db 238 TCGGAAATGGAGCGGCAGGTAAACGAGATCTGAAATGCTGTATGGGTATGTCGGCAA 297
QY 183 GCTTGTTCGTTCCACAATAA 203
Db 298 TCTTGTGTTTCCCGGTAAAA 318

RESULT 35
AAI67489
ID AAI67489 standard; DNA; 324 BP.
XX AC AAI67489;
XX DT 11-FEB-2002 (first entry)
XX DE Recombinant secretory leukocyte protease inhibitor (SLPI) DNA.
XX KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
XX KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
XX OS Homo sapiens.
XX OS US6291662-B1.
XX PN 18-SEP-2001.
XX PD

```

```

XX PF 22-SEP-1998; 98US-0158085.
XX PR 30-MAR-1987; 87US-0031846.
XX PR 06-AUG-1990; 90US-0563832.
XX PR 22-JUL-1994; 94US-0279056.
XX PR 05-DEC-1984; 84US-0678222.
XX PR 02-DEC-1985; 85US-0803471.
XX PR 29-JUL-1986; 86US-0890526.
XX PA (AMGE-) AMGEN INC.
XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX DR WPI; 2001-637974/73.
XX DR P-PSDB; AAG65996.
XX PT New DNA sequences, useful in recombinant DNA techniques for directing
XX PT the production of a serine protease inhibitor protein, e.g. leukocyte
XX PT elastase or trypsin -
XX PS Disclosure; Column 5; 37pp; English.
XX CC The invention relates to a DNA sequence encoding an analog of a mammalian
XX CC serine protease inhibitor protein. The analog comprises at least eight
XX CC cysteine residues and possesses serine protease inhibitor activity, and
XX CC at least one operational element consisting of a promoter, an operator,
XX CC or a terminator codon. The DNA is useful in recombinant DNA techniques
XX CC for directing the production of a serine protease inhibitor protein,
XX CC e.g. leukocyte elastase or trypsin. The present sequence represents the
XX CC coding strand of a DNA that directs the manufacture of a recombinant
XX CC secretory leukocyte protease inhibitor (SLPI).
XX SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match 15.6%; Score 32.2; DB 22; Length 324;
Best Local Similarity 51.8%; Pred. No. 0.63;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCAGGTTCTTGCTTATTTATCTTGGTTCGTCGCTATGTTAAACCCACCTAACCGT 122
Db 178 AAACCGGGTAAATGCCCGTAACCTATGCCAGTGTCTGATGCTGACCCCGCACTTC 237
QY 123 TGTTTGAAGGACACTGATGTCAGGTATCAAAAGTCTGTAAGGTTCTCGCGGTATG 182
Db 238 TCGGAAATGGAGCGGCAGGTAAACGAGATCTGAAATGCTGTATGGGTATGTCGGCAA 297
QY 183 GCTTGTTCGTTCCACAATAA 203
Db 298 TCTTGTGTTTCCCGGTAAAA 318

RESULT 36
AAI67489
ID AAI67489 standard; DNA; 420 BP.
XX AC AAI67489;
XX DT 19-APR-1999 (first entry)
XX DE DNA sequence of ompA-tc-met-SLPI.
XX KW Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
XX KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
XX KW acute leukemia; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN US5871956-A.
XX PD 16-FEB-1999.

```

XX PF 22-JUL-1994; 94US-0279056.
 XX PR 30-MAR-1987; 87US-0031846.
 XX PR 05-DEC-1984; 84US-0678822.
 XX PR 23-JUL-1986; 86US-0890526.
 XX PR 03-SEP-1986; 86US-0903471.
 XX PR 06-AUG-1990; 90US-0563832.
 XX PR 22-JUL-1994; 94US-0279056.
 XX PA (AMGE-) AMGEN INC.
 XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX DR WPI; 1999-166640/14.
 XX PT New DNA sequence encoding mammalian serine protease inhibitor -
 XX PT useful for recombinantly producing inhibitors with different
 XX PT specificities and treating diseases such as emphysema, arthritis,
 XX PT muscular dystrophy, and tumour invasion
 XX PS Example 3; Column 37-40; 37pp; English.
 XX CC The present invention describes a DNA sequence (A) encoding an analogue
 XX CC of a mammalian serine protease inhibitor (B). The DNA sequences and
 XX CC recombinant methods allow manufacture of a class of inhibitors of e.g.
 XX CC cathepsin G, elastase, and trypsin, with different specificities. The
 XX CC recombinant serine protease product can be directed to act
 XX CC intracellularly or extracellularly and is useful in treating conditions
 XX CC caused by a disturbance in the native protease/protease inhibitor
 XX CC balance such as emphysema, arthritis, glomerulonephritis, periodontitis,
 XX CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
 XX CC allows (B) to be recombinantly produced in sufficient quantities and
 XX CC purities so as to provide economical pharmaceutical compositions. (B) is
 XX CC resistant to heat, acid, and a variety of proteolytic enzymes, is
 XX CC thermodynamically stable in extracellular conditions, and exhibits a
 XX CC high degree of self assembly forming an active tertiary structure in the
 XX CC absence of biochemical stimuli. The present sequence represents the DNA
 XX CC sequence of ompA-tc-met-SLPI.
 XX SQ Sequence 420 BP; 123 A; 98 C; 112 G; 87 T; 0 other;
 Query Match 15.6%; Score 32.2; DB 20; Length 420;
 Best Local Similarity 51.8%; Pred. No. 0.69;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 63 AAGCCAGGTTCTTGTCTATTATCTTGGATTTCGTTGGCTATGTAAACCCACCTAACCGT 122
 DB 278 AAACCCGGTAAATGCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCCGCAACTTC 337
 QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTTCCTCGCGTATG 182
 DB 338 TGCAGAAATGACGGCCAGTGTAAACGAGATCTGAATGCTGTATGGTATGTCGGCAAA 397
 QY 183 GCTTGTTCGTTCCACAATAA 203
 DB 398 TCTTGTGTTTCCCGGTAATAA 418
 RESULT 37
 AAC97624
 ID AAC97624 standard; DNA; 436 BP.
 XX AC AAC97624;
 XX DT 27-FEB-2001 (first entry)
 XX DE DNA encoding ompA-tc-met-SLPI.
 XX KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.

XX OS Synthetic.
 XX PN US6132990-A.
 XX PD 17-OCT-2000.
 XX PF 07-JUN-1991; 91US-0712354.
 XX PR 03-JAN-1989; 89US-0293042.
 XX PR 06-DEC-1984; 84US-0678822.
 XX PR 02-DEC-1985; 85US-0803471.
 XX PR 29-JUL-1986; 86US-0890526.
 XX PR 30-MAR-1987; 87US-0031846.
 XX PR 04-AUG-1987; 87US-0082962.
 XX PA (AMGE-) AMGEN BOULDER INC.
 XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX DR WPI; 2000-678667/66.
 XX PT New serine protease inhibitors and DNA sequences for treating a
 XX PT protease-mediated condition or tissue destruction e.g. emphysema or
 XX PS tumour invasion and for recombinant production of inhibitors -
 XX PS Example 3; Column 39-40; 47pp; English.
 XX CC This invention relates to new purified and isolated mammalian serine
 XX CC protease inhibitor proteins which comprise at least 8 cysteine residues
 XX CC and no more than 107 amino acids. The protease inhibitors are capable of
 XX CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534
 XX CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 XX CC protease inhibitors of the invention, and include oligonucleotide
 XX CC sequences used in the isolation and characterisation of the proteins.
 XX CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 XX CC construction of DNA encoding the protease inhibitors. Peptide sequences
 XX CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 XX CC various other peptides used in the isolation of the protease inhibitors.
 XX CC The serine protease inhibitors have cytostatic and anti-inflammatory activity.
 XX CC The serine protease inhibitor protein is useful for treating a
 XX CC protease-mediated condition, which includes protease mediated tissue
 XX CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 XX CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 XX CC for modulating protease activity. The DNAs are useful for producing the
 XX CC proteins, especially by recombinant methods.
 XX SQ Sequence 436 BP; 128 A; 102 C; 116 G; 90 T; 0 other;
 Query Match 15.6%; Score 32.2; DB 21; Length 436;
 Best Local Similarity 51.8%; Pred. No. 0.7;
 Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 63 AAGCCAGGTTCTTGTCTATTATCTTGGATTTCGTTGGCTATGTAAACCCACCTAACCGT 122
 DB 278 AAACCCGGTAAATGCCCGTAAACCTATGGCCAGTGTCTGATGCTGAACCCCGCAACTTC 337
 QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTTCCTCGCGTATG 182
 DB 338 TGCAGAAATGACGGCCAGTGTAAACGAGATCTGAATGCTGTATGGTATGTCGGCAAA 397
 QY 183 GCTTGTTCGTTCCACAATAA 203
 DB 398 TCTTGTGTTTCCCGGTAATAA 418
 RESULT 38
 AAI67583
 ID AAI67583 standard; DNA; 436 BP.
 XX AC AAI67583;
 XX DT 11-FEB-2002 (first entry)

```

XX DE DNA sequence coding for ompA-tc-met-SLPI.
XX KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
XX KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
XX OS Homo sapiens.
XX PN US6291662-B1.
XX PD 18-SEP-2001.
XX PF 22-SEP-1998; 98US-0158085.
XX PR 30-MAR-1987; 87US-0031846.
XX PR 06-AUG-1990; 90US-0563832.
XX PR 22-JUL-1994; 94US-0279056.
XX PR 05-DEC-1984; 84US-0678222.
XX PR 02-DEC-1985; 85US-0803471.
XX PR 29-JUL-1986; 86US-0890526.
XX PA (AMGE-) AMGEN INC.
XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX DR WPI; 2001-637974/73.
XX KW New DNA sequences, useful in recombinant DNA techniques for directing
XX PT the production of a serine protease inhibitor protein, e.g. leukocyte
XX PT elastase or trypsin.
XX PS Example 3; Column 37-38; 37pp; English.
XX CC The invention relates to a DNA sequence encoding an analog of a mammalian
XX CC serine protease inhibitor protein. The analog comprises at least eight
XX CC cysteine residues and possesses serine protease inhibitor activity, and
XX CC at least one operational element consisting of a promoter, an operator,
XX CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
XX CC or a terminator codon. The DNA is useful in recombinant DNA techniques
XX CC for directing the production of a serine protease inhibitor protein,
XX CC e.g. leukocyte elastase or trypsin. The present sequence represents the
XX CC DNA sequence coding for ompA-tc-met-secretory leukocyte protease
XX CC inhibitor (SLPI).
XX SQ Sequence 436 BP; 128 A; 102 C; 116 G; 90 T; 0 other;

Query Match 15.6%; Score 32.2; DB 22; Length 436;
Best Local Similarity 51.8%; Pred. No. 0.7;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGCTATTATCTTGATTCGTTGGCGTATGTTAAACCCACCTTAACCGT 122
Db 278 AAACCGGGTAAATGCCCGGTAACTTATGCCAGTGCTGATGCTGAACCCCGCGAACTTC 337
QY 123 TGTTTGAAGGACACTGATTTGTCAGGTATCAAAAGTGTGGAAGTTCCTGCGGTATG 182
Db 338 TGGCAATGGACGGCCAGGTGTAACAGAGATCTGAAATGCTGTATGGTGTGCGGCAA 397
QY 183 GCTTGTTCGTTCCACAATAA 203
Db 398 TCTTGTGTTCCCGGTAAAA 418

RESULT 39
AAI16271
ID AAX16271 standard; DNA; 460 BP.
XX AC AAX16271;
XX 19-APR-1999 (first entry)
XX 19-APR-1999 (first entry)
XX DNA sequence of ompA-SLPI.
XX

```

```

KW Serine protease inhibitor; human; emphysema; arthritis; periodontitis;
KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW acute leukemia; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN US5871956-A.
XX PD 16-FEB-1999.
XX PF 22-JUL-1994; 94US-0279056.
XX PR 30-MAR-1987; 87US-0031846.
XX PR 05-DEC-1984; 84US-0678822.
XX PR 29-JUL-1986; 86US-0890526.
XX PR 03-SEP-1986; 86US-0903471.
XX PR 06-AUG-1990; 90US-0563832.
XX PR 22-JUL-1994; 94US-0279056.
XX PA (AMGE-) AMGEN INC.
XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX DR WPI; 1999-166640/14.
XX KW New DNA sequence encoding mammalian serine protease inhibitor -
XX PT useful for recombinantly producing inhibitors with different
XX PT specificities and treating diseases such as emphysema, arthritis,
XX PT muscular dystrophy, and tumour invasion
XX PS Example 3; Column 37-38; 37pp; English.
XX CC The present invention describes a DNA sequence (A) encoding an analogue
XX CC of a mammalian serine protease inhibitor (B). The DNA sequences and
XX CC recombinant methods allow manufacture of a class of inhibitors of e.g.
XX CC cathepsin G, elastase, and trypsin, with different specificities. The
XX CC recombinant serine protease product can be directed to act
XX CC intracellularly or extracellularly and is useful in treating conditions
XX CC caused by a disturbance in the native protease/protease inhibitor
XX CC balance such as emphysema, arthritis, glomerulonephritis, periodontitis,
XX CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
XX CC allows (B) to be recombinantly produced in sufficient quantities and
XX CC purities so as to provide economical pharmaceutical compositions. (B) is
XX CC resistant to heat, acid, and a variety of proteolytic enzymes, is
XX CC thermodynamically stable in extracellular conditions, and exhibits a
XX CC high degree of self assembly forming an active tertiary structure in the
XX CC absence of biochemical stimuli. The present sequence represents the DNA
XX CC sequence of ompA-SLPI.
XX SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Query Match 15.6%; Score 32.2; DB 20; Length 460;
Best Local Similarity 51.8%; Pred. No. 0.72;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGCTATTATCTTGATTCGTTGGCGTATGTTAAACCCACCTTAACCGT 122
Db 302 AAACCGGGTAAATGCCCGGTAACTTATGCCAGTGCTGATGCTGAACCCCGCGAACTTC 361
QY 123 TGTTTGAAGGACACTGATTTGTCAGGTATCAAAAGTGTGGAAGTTCCTGCGGTATG 182
Db 362 TGGCAATGGACGGCCAGGTGTAACAGAGATCTGAAATGCTGTATGGTGTGCGGCAA 421
QY 183 GCTTGTTCGTTCCACAATAA 203
Db 422 TCTTGTGTTCCCGGTAAAA 442

RESULT 40
AAC97622
ID AAC97622 standard; DNA; 460 BP.
XX

```

AC AAC97622;
 XX
 DT 27-FEB-2001 (first entry)
 DE
 DE DNA encoding OmpA secretory leukocyte protease inhibitor.
 XX
 KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
 KW protease mediated tissue destruction; emphysema; glomerulonephritis;
 KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
 KW elastase; ds.
 XX
 OS Synthetic.
 XX
 XX US6132990-A.
 PN
 XX 17-OCT-2000.
 PD
 XX
 XX 07-JUN-1991; 91US-0712354.
 PF
 XX
 PR 03-JAN-1989; 89US-0293042.
 PR 06-DEC-1984; 84US-0678822.
 PR 02-DEC-1985; 85US-0803471.
 PR 29-JUL-1986; 86US-0890526.
 PR 30-MAR-1987; 87US-0031846.
 PR 04-AUG-1987; 87US-0082962.
 XX
 PA (AMGE-) AMGEN BOULDER INC.
 XX
 PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX
 DR WPI; 2000-678667/66.
 XX
 XX New serine protease inhibitors and DNA sequences for treating a
 PT protease-mediated condition or tissue destruction e.g. emphysema or
 PT tumor invasion and for recombinant production of inhibitors
 XX
 PS Example 3; Column 37-38; 47pp; English.
 XX
 CC This invention relates to new purified and isolated mammalian serine
 CC protease inhibitor proteins which comprise at least 8 cysteine residues
 CC and no more than 107 amino acids. The protease inhibitors are capable of
 CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
 CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
 CC protease inhibitors of the invention, and include oligonucleotide
 CC sequences used in the isolation and characterisation of the proteins.
 CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
 CC construction of DNA encoding the protease inhibitors. Peptide sequences
 CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
 CC various other peptides used in the isolation of the protease inhibitors.
 CC The protease inhibitors have cytostatic and anti-inflammatory activity.
 CC The serine protease inhibitor protein is useful for treating a
 CC protease-mediated condition, which includes protease mediated tissue
 CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
 CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
 CC for modulating protease activity. The DNAs are useful for producing the
 CC proteins, especially by recombinant methods.
 XX
 SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;
 Query Match 15.6%; Score 32.2; DB 21; Length 460;
 Best Local Similarity 51.8%; Pred. No. 0.72; Indels 0; Gaps 0;
 Matches 73; Conservative 0; Mismatches 68;
 QY 63 AAGCCAGGTTCTTGCTTATTTATCTGTTGGCTATGTTAAACCCCACTAACCGT 122
 DB 302 AAACCGGGTAAATGCCCGGTAACTATGCGCAGTGCTGATGCTGAACCCCGCACTTC 361
 QY 123 TGTTTGAAGGACACTGATGTCTCAGGTATCAAAAAGTGTGTAAGGTTCTTCGCGGTATG 182
 DB 362 TCGAAATGACGGCCAGGTGTAACAGAGATCTGAAATGCTGTATGGGTATGTGCGCAAA 421
 QY 183 GCTTGTTCGTTCCACAATAA 203

Db 422 TCTTGTGTTCCTCCCGGTAAAA 442
 RESULT 41
 AAI67582
 ID AAI67582 standard; DNA; 460 BP.
 XX
 AC AAI67582;
 XX
 DT 11-FEB-2002 (first entry)
 XX
 DE DNA sequence coding for ompA SLPI.
 XX
 KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
 KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
 XX
 OS Homo sapiens.
 XX
 PN US6291662-B1.
 XX
 PD 18-SEP-2001.
 XX
 PF 22-SEP-1998; 98US-0158085.
 XX
 PR 30-MAR-1987; 87US-0031846.
 PR 06-AUG-1990; 90US-0563832.
 PR 22-JUL-1994; 94US-0279056.
 PR 05-DEC-1984; 84US-0678222.
 PR 02-DEC-1985; 85US-0803471.
 PR 29-JUL-1986; 86US-0890526.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
 XX
 DR WPI; 2001-637974/73.
 XX
 XX New DNA sequences, useful in recombinant DNA techniques for directing
 PT the production of a serine protease inhibitor protein, e.g. leukocyte
 PT elastase or trypsin
 XX
 PS Example 3; Column 36; 37pp; English.
 XX
 CC The invention relates to a DNA sequence encoding an analog of a mammalian
 CC serine protease inhibitor protein. The analog comprises at least eight
 CC cysteine residues and possesses serine protease inhibitor activity, and
 CC at least one operational element consisting of a promoter, an operator,
 CC or a terminator codon. The DNA is useful in recombinant DNA techniques
 CC for directing the production of a serine protease inhibitor protein,
 CC e.g. leukocyte elastase or trypsin. The present sequence represents the
 CC DNA sequence coding for ompA-secretory leukocyte protease inhibitor
 CC (SLPI).
 XX
 SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;
 Query Match 15.6%; Score 32.2; DB 22; Length 460;
 Best Local Similarity 51.8%; Pred. No. 0.72; Indels 0; Gaps 0;
 Matches 73; Conservative 0; Mismatches 68;
 QY 63 AAGCCAGGTTCTTGCTTATTTATCTGTTGGCTATGTTAAACCCCACTAACCGT 122
 DB 302 AAACCGGGTAAATGCCCGGTAACTATGCGCAGTGCTGATGCTGAACCCCGCACTTC 361
 QY 123 TGTTTGAAGGACACTGATGTCTCAGGTATCAAAAAGTGTGTAAGGTTCTTCGCGGTATG 182
 DB 362 TCGAAATGACGGCCAGGTGTAACAGAGATCTGAAATGCTGTATGGGTATGTGCGCAAA 421
 QY 183 GCTTGTTCGTTCCACAATAA 203
 Db 422 TCTTGTGTTCCTCCCGGTAAAA 442

```

RESULT 42
AAS51617/c
ID AAS51617 standard; DNA; 546 BP.
XX
AC AAS51617;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #34.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-207727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
FA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU33758.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 4199; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acid sequence is also
CC used to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 546 BP; 205 A; 73 C; 122 G; 146 T; 0 other;
Query Match 15.6%; Score 32.2; DB 23; Length 546;
Best Local Similarity 48.1%; Pred. No. 0.76;
Matches 91; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 2 ATTGAGCTCGGTACCACTACCTGCGTATGCTCAAGAACAGGTTAAGCTCTGCTGTAC 61
DB 312 ATTGGCTTAGACCCCTGCACCTGCAGAACCTACAAACAGCAAGCCCTGCTGTATTCT 253
QY 62 TAAGCCAGGTTCTTGCTCTATTATCTTGATTCGTTGGCGTATGTTAAACCCCACTAACCG 121

```

```

Db 252 TACCACATACCATTGATTCATCTGTCATGATTAATTCACATAAAACATATCCAGGGAATGT 193
QY 122 TTGTTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGCTGTGAAGGTTCTCTCGGTAT 181
Db 192 TTTTAAACGGTGTGTTTAGCTTTACCATCTTTAACTTGAGTTCTTCTTCTTCCGGTAT 133
QY 182 GGCTTGTTT 190
Db 132 GACTACTCT 124
RESULT 43
AAS54717/c
ID AAS54717 standard; DNA; 549 BP.
XX
AC AAS54717;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1029.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-207727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU36858.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 8354; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acid sequence is also
CC used to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

```

XX
SQ Sequence 549 BP; 207 A; 73 C; 122 G; 147 T; 0 other;
Query Match 15.6%; Score 32.2; DB 23; Length 549;
Best Local Similarity 48.1%; Pred. No. 0.77; 98; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
OY 2 ATTGAGCTCGGTACCATACCTGCATATGCTCAAGAACCGATTAAAGGTCTGTGTCTAC 61
DB 312 ATTGGCTTAGACCTGCGACCTGCAGAACCTACAAACAGTAACGCCCTGGTGTATTTCT 253
OY 62 TAAGCCAGGTCTTGTGCTTATTTATCTTGATTCGTTGCGGTATGTTAAACCCACCTAACCG 121
DB 252 TACCACATACCATGATTCATCTGTCAATGTAATTCCTAAACATATCCAGGGAATGT 193
OY 122 TTGTTTGAAGGACACTGATGTCAGGTATCAAAAGTCTGTAAGGTTCCTGCGGTAT 181
DB 192 TTTTAAACGGTGTGTTTGTAGCTTACCATCTTTAACTTGAGTTCTTCTTCCGGTAT 133
OY 182 GGCTGTTT 190
DB 132 GACTACTCT 124
RESULT 44
AAK65489
ID AAK65489 standard; DNA; 16605 BP.
XX
AC AAK65489;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20301.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
FN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.

PR	08-NOV-2000;	2000US-0246610.	
PR	08-NOV-2000;	2000US-0246611.	
PR	08-NOV-2000;	2000US-0246613.	
PR	17-NOV-2000;	2000US-0249207.	
PR	17-NOV-2000;	2000US-0249208.	
PR	17-NOV-2000;	2000US-0249209.	
PR	17-NOV-2000;	2000US-0249210.	
PR	17-NOV-2000;	2000US-0249211.	
PR	17-NOV-2000;	2000US-0249212.	
PR	17-NOV-2000;	2000US-0249213.	
PR	17-NOV-2000;	2000US-0249215.	
PR	17-NOV-2000;	2000US-0249216.	
PR	17-NOV-2000;	2000US-0249217.	
PR	17-NOV-2000;	2000US-0249218.	
PR	17-NOV-2000;	2000US-0249244.	
PR	17-NOV-2000;	2000US-0249245.	
PR	17-NOV-2000;	2000US-0249264.	
PR	17-NOV-2000;	2000US-0249265.	
PR	17-NOV-2000;	2000US-0249297.	
PR	17-NOV-2000;	2000US-0249299.	
PR	17-NOV-2000;	2000US-0249300.	
PR	01-DEC-2000;	2000US-0250160.	
PR	01-DEC-2000;	2000US-0250391.	
PR	05-DEC-2000;	2000US-0251030.	
PR	05-DEC-2000;	2000US-0251988.	
PR	05-DEC-2000;	2000US-0256719.	
PR	06-DEC-2000;	2000US-0251479.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251868.	
PR	08-DEC-2000;	2000US-0251869.	
PR	08-DEC-2000;	2000US-0251989.	
PR	08-DEC-2000;	2000US-0251990.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0259678.	
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Barash SC, Ruben SM;		
XX	WPI; 2001-483426/52.		
DR			
XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,		
PT	useful for preventing, diagnosing and/or treating cancers and		
PT	metastasis		
XX			
PS	Disclosure; SEQ ID NO 20301; 3071pp + Sequence Listing; English.		
XX			
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)		
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic		
CC	activity, and can be used in gene therapy and vaccine production. (I)		
CC	proteins, and polynucleotides may be used in the prevention, diagnosis and		
CC	treatment of diseases associated with inappropriate (I) expression. For		
CC	example, they may be used to treat disorders associated with decreased		
CC	expression by rectifying mutations or deletions in a patient's genome		
CC	that affect the activity of (I) by expressing inactive proteins or to		
CC	supplement the patient's own production of (I). Additionally, (I)		
CC	polynucleotides may be used to produce the secreted (I), by inserting		
CC	the nucleic acids into a host cell and culturing the cell to express the		
CC	protein. (I) treats and polynucleotides may be used to prevent,		
CC	diagnose and treat immune/haematopoietic-related diseases, especially		
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703		
CC	to AAK87694 represent human immune/haematopoietic antigen genomic		
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169		
CC	represent sequences used in the exemplification of the present invention.		
XX			
XX	Sequence 16605 BP; 4009 A; 3696 C; 4212 G; 4688 T; 0 other;		
SQ			
	Query Match	15.3%;	Score 31.6; DB 22; Length 16605;
	Best Local Similarity	52.2%;	Pred. No. 4.4;
	Matches	70; Conservative	0; Mismatches 64; Indels 0; Gaps 0;
QY	55	TGCTACTAGCAGGTTCTTGTCTTATTTCTGAGCCATTTATTTCTGACCTGGGCGAGGATCACTCACT	4063

Db	4004	TTTCTATTATGTTCTTCTTTCTTAGGCCATTTATTTCTGACCTGGGCGAGGATCACTCACT	4063
QY	115	CTAACCGTTGTTTGAAGGACACTGATTTGTCAGGTATCAAAAAGTGTGTGAAGTTTCT	174
Db	4064	TGGGCGCTTATAGCTGGGACACTGACGCTCAAAATACAGAGCTGCTTGAATGGGTATT	4123
QY	175	GCGGTATGGCTTGT	188
Db	4124	GTAATATGATGGT	4137
RESULT	45		
AAK87049			
ID	AAK87049	standard; DNA; 16605 BP.	
XX			
AC	AAK87049;		
XX			
DT	07-NOV-2001	(first entry)	
XX			
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41861.		
XX			
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
KW	cytostatic; gene therapy; vaccine; metastasis; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200157182-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01354.		
XX			
XX	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0226279.		
PR	22-AUG-2000; 2000US-0226681.		
PR	22-AUG-2000; 2000US-0226686.		
PR	22-AUG-2000; 2000US-0227182.		
PR	23-AUG-2000; 2000US-0227009.		
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0229287.		
PR	01-SEP-2000; 2000US-0229343.		
PR	01-SEP-2000; 2000US-0229344.		
PR	01-SEP-2000; 2000US-0229345.		

PR	05-SEP-2000;	2000US-02295009
PR	05-SEP-2000;	2000US-02295010
PR	06-SEP-2000;	2000US-02295437
PR	06-SEP-2000;	2000US-02304337
PR	06-SEP-2000;	2000US-02304338
PR	08-SEP-2000;	2000US-02312422
PR	08-SEP-2000;	2000US-02312423
PR	08-SEP-2000;	2000US-02312424
PR	08-SEP-2000;	2000US-02314133
PR	08-SEP-2000;	2000US-02314134
PR	08-SEP-2000;	2000US-02314141
PR	08-SEP-2000;	2000US-02320800
PR	08-SEP-2000;	2000US-02320811
PR	12-SEP-2000;	2000US-02331968
PR	14-SEP-2000;	2000US-02332397
PR	14-SEP-2000;	2000US-02332398
PR	14-SEP-2000;	2000US-02332399
PR	14-SEP-2000;	2000US-02332400
PR	14-SEP-2000;	2000US-02334011
PR	14-SEP-2000;	2000US-02330631
PR	14-SEP-2000;	2000US-02330632
PR	14-SEP-2000;	2000US-02330633
PR	21-SEP-2000;	2000US-02334623
PR	21-SEP-2000;	2000US-02342274
PR	25-SEP-2000;	2000US-02349977
PR	25-SEP-2000;	2000US-02349978
PR	25-SEP-2000;	2000US-02349980
PR	26-SEP-2000;	2000US-02349884
PR	27-SEP-2000;	2000US-02358344
PR	27-SEP-2000;	2000US-02358346
PR	29-SEP-2000;	2000US-02363127
PR	29-SEP-2000;	2000US-02363127
PR	29-SEP-2000;	2000US-02363167
PR	29-SEP-2000;	2000US-02363668
PR	29-SEP-2000;	2000US-02363669
PR	29-SEP-2000;	2000US-02363700
PR	02-OCT-2000;	2000US-0236802
PR	02-OCT-2000;	2000US-02370377
PR	02-OCT-2000;	2000US-02370388
PR	02-OCT-2000;	2000US-02370399
PR	02-OCT-2000;	2000US-02370399
PR	13-OCT-2000;	2000US-02399335
PR	13-OCT-2000;	2000US-02399337
PR	20-OCT-2000;	2000US-02409600
PR	20-OCT-2000;	2000US-02412221
PR	20-OCT-2000;	2000US-02417855
PR	20-OCT-2000;	2000US-02417866
PR	20-OCT-2000;	2000US-02417877
PR	20-OCT-2000;	2000US-02464575
PR	20-OCT-2000;	2000US-02464576
PR	20-OCT-2000;	2000US-02418089
PR	20-OCT-2000;	2000US-02418090
PR	01-NOV-2000;	2000US-02446177
PR	08-NOV-2000;	2000US-02464774
PR	08-NOV-2000;	2000US-02464775
PR	08-NOV-2000;	2000US-02465225
PR	08-NOV-2000;	2000US-02465226
PR	08-NOV-2000;	2000US-02465227
PR	08-NOV-2000;	2000US-02465228
PR	08-NOV-2000;	2000US-02465232
PR	08-NOV-2000;	2000US-02465609
PR	08-NOV-2000;	2000US-02466110
PR	08-NOV-2000;	2000US-02466111
PR	08-NOV-2000;	2000US-02466113
PR	17-NOV-2000;	2000US-02492007
PR	17-NOV-2000;	2000US-02492008
PR	17-NOV-2000;	2000US-02492010
PR	17-NOV-2000;	2000US-02492019
PR	17-NOV-2000;	2000US-02492111
PR	17-NOV-2000;	2000US-02492122
PR	17-NOV-2000;	2000US-02492134
PR	17-NOV-2000;	2000US-02492145

PR	17-NOV-2000;	2000US-0249216
PR	17-NOV-2000;	2000US-0249217
PR	17-NOV-2000;	2000US-0249218
PR	17-NOV-2000;	2000US-0249219
PR	17-NOV-2000;	2000US-0249244
PR	17-NOV-2000;	2000US-0249245
PR	17-NOV-2000;	2000US-0249264
PR	17-NOV-2000;	2000US-0249265
PR	17-NOV-2000;	2000US-0249297
PR	17-NOV-2000;	2000US-0249299
PR	17-NOV-2000;	2000US-0249300
PR	01-DEC-2000;	2000US-0249300
PR	01-DEC-2000;	2000US-0250160
PR	01-DEC-2000;	2000US-0250391
PR	05-DEC-2000;	2000US-0251030
PR	05-DEC-2000;	2000US-0251988
PR	05-DEC-2000;	2000US-0256719
PR	06-DEC-2000;	2000US-0251479
PR	08-DEC-2000;	2000US-0251856
PR	08-DEC-2000;	2000US-0251968
PR	08-DEC-2000;	2000US-0251969
PR	08-DEC-2000;	2000US-0251989
PR	08-DEC-2000;	2000US-0251990
PR	11-DEC-2000;	2000US-0254097
PR	05-JAN-2001;	2001US-0254967

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 41861; 3071pp + Sequence Listing: English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM821169 represent sequences used in the exemplification of the present invention.

Sequence 16605 BP; 4009 A; 3696 C; 4212 G; 4688 T; 0 other;

Query Match 15.3%; Score 31.6; DB 22; Length 16605;
Best Local Similarity 52.2%; Pred. No. 4.4;
Matches 70: Conservative 0; Mismatches 64; Indels 0; Gaps 0

QY		55	TGCTACTAAGCAGGTTC TTGGTCCATTTATCTTGATT CGTTGC GTATGTTAAACCCAC	114
D6		4004	TTTCTATATG TCTCTT AGSCCATTTT CTGCCCT GGGCAGG ATCACT CACT	4063

QY 115 CTAACCGTTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTT 174

Db 4064 TGGGGCCTTATAGCTGGGACACTGACGCTCAAAATACCAGGAGCTGCTGGAATGGGTATT 4123

Qy 175 GCGGTATGGCTTGT 188

Db 4124 GTAAATGTATGGT 4137

```

RESULT 46
ABV53540/c
ID ABV53540 standard; cDNA; 612 BP.
XX
XX
AC ABV53540;
XX
XX
DT 17-SEP-2002 (first entry)
XX
XX
DE Human prostate expression marker cDNA 53531.
XX
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX
OS Homo sapiens.
XX
XX
FN WO200160860-A2.
XX
XX
PD 23-AUG-2001.
XX
XX
FF 20-FEB-2001; 2001WO-US05171.
XX
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX
DR WPI; 2001-662795/76.
XX
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
XX
PS Claim 1; Page 10369; 11750pp; English.
XX
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX
SQ Sequence 612 BP; 241 A; 100 C; 102 G; 169 T; 0 other;

Query Match 14.8%; Score 30.4; DB 23; Length 612;
Best Local Similarity 63.9%; Pred. No. 3.3;
Matches 46; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 60 ACTAGCCAGGTTCTTGCTATTTATTTGATTCGTTGGGTATGTTAAACCCACCTTAAC 119
Db 306 AATAAGAGTGATTCCTTGTTATGATTCCTGATCTGTCGATCTCTAAATTCATTAAT 247
QY 120 CGTGTGTTGAAG 131
Db 246 AGTTACATGAGG 235

RESULT 47
ABA18664/c

```

```

ID ABA18664 standard; DNA; 3496 BP.
XX
XX
AC ABA18664;
XX
XX
DT 23-JAN-2002 (first entry)
XX
XX
DE Human nervous system related polynucleotide SEQ ID NO 10995.
XX
XX
KW Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
XX
XX
OS Homo sapiens.
XX
XX
FN WO200159063-A2.
XX
XX
PD 16-AUG-2001.
XX
XX
FF 17-JAN-2001; 2001WO-US01334.
XX
XX
PR 31-JAN-2000; 2000US-0179065.
XX
PR 04-FEB-2000; 2000US-0180628.
XX
PR 24-FEB-2000; 2000US-0184664.
XX
PR 02-MAR-2000; 2000US-0186350.
XX
PR 16-MAR-2000; 2000US-0189874.
XX
PR 17-MAR-2000; 2000US-0190076.
XX
PR 18-APR-2000; 2000US-0198123.
XX
PR 19-MAY-2000; 2000US-0205151.
XX
PR 07-JUN-2000; 2000US-0209467.
XX
PR 28-JUN-2000; 2000US-0214886.
XX
PR 30-JUN-2000; 2000US-0215135.
XX
PR 07-JUL-2000; 2000US-0216647.
XX
PR 07-JUL-2000; 2000US-0216880.
XX
PR 11-JUL-2000; 2000US-0217487.
XX
PR 11-JUL-2000; 2000US-0217496.
XX
PR 14-JUL-2000; 2000US-0218290.
XX
PR 26-JUL-2000; 2000US-0220963.
XX
PR 26-JUL-2000; 2000US-0220964.
XX
PR 14-AUG-2000; 2000US-0224518.
XX
PR 14-AUG-2000; 2000US-0224519.
XX
PR 14-AUG-2000; 2000US-0235213.
XX
PR 14-AUG-2000; 2000US-0235214.
XX
PR 14-AUG-2000; 2000US-0235266.
XX
PR 14-AUG-2000; 2000US-0235267.
XX
PR 14-AUG-2000; 2000US-0235268.
XX
PR 14-AUG-2000; 2000US-0235270.
XX
PR 14-AUG-2000; 2000US-0235447.
XX
PR 14-AUG-2000; 2000US-0235757.
XX
PR 14-AUG-2000; 2000US-0235758.
XX
PR 14-AUG-2000; 2000US-0235759.
XX
PR 18-AUG-2000; 2000US-0226279.
XX
PR 22-AUG-2000; 2000US-0226681.
XX
PR 22-AUG-2000; 2000US-0226868.
XX
PR 22-AUG-2000; 2000US-0227182.
XX
PR 23-AUG-2000; 2000US-0227009.
XX
PR 30-AUG-2000; 2000US-0228924.
XX
PR 01-SEP-2000; 2000US-0229287.
XX
PR 01-SEP-2000; 2000US-0229343.
XX
PR 01-SEP-2000; 2000US-0229344.
XX
PR 01-SEP-2000; 2000US-0229345.
XX
PR 05-SEP-2000; 2000US-0229509.
XX
PR 05-SEP-2000; 2000US-0229513.
XX
PR 06-SEP-2000; 2000US-0230437.
XX
PR 06-SEP-2000; 2000US-0230438.
XX
PR 08-SEP-2000; 2000US-0231242.
XX
PR 08-SEP-2000; 2000US-0231243.
XX
PR 08-SEP-2000; 2000US-0231244.
XX
PR 08-SEP-2000; 2000US-0231413.
XX
PR 08-SEP-2000; 2000US-0231414.

```

PR	08-SEP-2000;	2000US-02320800;
PR	08-SEP-2000;	2000US-02320811;
PR	12-SEP-2000;	2000US-02321968;
PR	14-SEP-2000;	2000US-02323397;
PR	14-SEP-2000;	2000US-02323398;
PR	14-SEP-2000;	2000US-02323399;
PR	14-SEP-2000;	2000US-02324000;
PR	14-SEP-2000;	2000US-02324001;
PR	14-SEP-2000;	2000US-02330663;
PR	14-SEP-2000;	2000US-02330664;
PR	14-SEP-2000;	2000US-02330665;
PR	21-SEP-2000;	2000US-02342231;
PR	21-SEP-2000;	2000US-02342274;
PR	25-SEP-2000;	2000US-02343997;
PR	25-SEP-2000;	2000US-02343998;
PR	26-SEP-2000;	2000US-02354884;
PR	27-SEP-2000;	2000US-02358834;
PR	27-SEP-2000;	2000US-02358836;
PR	29-SEP-2000;	2000US-02363237;
PR	29-SEP-2000;	2000US-02363267;
PR	29-SEP-2000;	2000US-02363668;
PR	29-SEP-2000;	2000US-02363669;
PR	02-OCT-2000;	2000US-02363707;
PR	02-OCT-2000;	2000US-02368802;
PR	02-OCT-2000;	2000US-02370037;
PR	02-OCT-2000;	2000US-02370038;
PR	02-OCT-2000;	2000US-02370039;
PR	02-OCT-2000;	2000US-02370040;
PR	13-OCT-2000;	2000US-02399935;
PR	20-OCT-2000;	2000US-02409660;
PR	20-OCT-2000;	2000US-02417885;
PR	20-OCT-2000;	2000US-02417886;
PR	20-OCT-2000;	2000US-02417887;
PR	20-OCT-2000;	2000US-02418008;
PR	20-OCT-2000;	2000US-02418009;
PR	20-OCT-2000;	2000US-02418236;
PR	01-NOV-2000;	2000US-02422211;
PR	01-NOV-2000;	2000US-02424617;
PR	08-NOV-2000;	2000US-02464674;
PR	08-NOV-2000;	2000US-02464675;
PR	08-NOV-2000;	2000US-02464676;
PR	08-NOV-2000;	2000US-02464677;
PR	08-NOV-2000;	2000US-02464678;
PR	08-NOV-2000;	2000US-02465223;
PR	08-NOV-2000;	2000US-02465224;
PR	08-NOV-2000;	2000US-02465225;
PR	08-NOV-2000;	2000US-02465226;
PR	08-NOV-2000;	2000US-02465227;
PR	08-NOV-2000;	2000US-02465228;
PR	08-NOV-2000;	2000US-02465229;
PR	08-NOV-2000;	2000US-02465230;
PR	08-NOV-2000;	2000US-02466110;
PR	08-NOV-2000;	2000US-02466111;
PR	08-NOV-2000;	2000US-02466113;
PR	17-NOV-2000;	2000US-02492077;
PR	17-NOV-2000;	2000US-02492078;
PR	17-NOV-2000;	2000US-02492114;
PR	17-NOV-2000;	2000US-02492115;
PR	17-NOV-2000;	2000US-02492116;
PR	17-NOV-2000;	2000US-02492117;
PR	17-NOV-2000;	2000US-02492118;
PR	17-NOV-2000;	2000US-02492444;
PR	17-NOV-2000;	2000US-02492445;
PR	17-NOV-2000;	2000US-02492464;
PR	17-NOV-2000;	2000US-02492465;
PR	17-NOV-2000;	2000US-02492997;
PR	17-NOV-2000;	2000US-02492999;

PR	17-NOV-2000;	2000US-0249300
PR	01-DEC-2000;	2000US-0250391
PR	01-DEC-2000;	2000US-0251160
PR	03-DEC-2000;	2000US-0251030
PR	05-DEC-2000;	2000US-0251988
PR	05-DEC-2000;	2000US-0256719
PR	06-DEC-2000;	2000US-0251479
PR	08-DEC-2000;	2000US-0251856
PR	08-DEC-2000;	2000US-0251868
PR	08-DEC-2000;	2000US-0251869
PR	08-DEC-2000;	2000US-0251989
PR	08-DEC-2000;	2000US-0251990
PR	11-DEC-2000;	2000US-0254097
PR	05-JAN-2001;	2001US-0259678

Query Match 14.8%; Score 30.4; DB 22; Length 3496;
Best Local Similarity 49.4%; Pred. No. 6.2;

QY 30 GCTCAAGAACGATTAAAGGTCCCTGTGTCTACTTAAGCAGGTTCTTGTCCTATTAATCTTG 89
 ||| | ||||| || || || || || || ||
Dδ 1085 GTTCTGGTTCAGGTC AAGGGCAAAAATGATGATA TTTCAATTTTCCTGTTCTACATGCTTA 1026
 ||| | ||||| || || || || || || ||

QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGCACACTGATTGTCAGGT 149
||| | | | | | | | | | | | | | |
Db 1025 ATAGTCAGTGAAATATTTGGCCCAAGAGGCCATGCTATGTTGGATTTTAATACAAAAGGT 966

QY
150 ATCAAAAAGTCTGTGAAGGTTCTCGGGTATGGCTTGTT 189

Dδ
965 AAATAAAGCTCTGCTGATGTCAGGCCAAAAATTATGCT 926

RESULT 48
ABQ88179
ID ABQ88179 standard; cDNA; 160771 BP.
XX
XX
XX
XX
XX
XX
DT 18-SEP-2002 (first entry)

DE Human osteoblast differentiation related cDNA SEQ ID NO 86.
 XX Human; osteoblast; stem cell differentiation; bone tissue deposition;
 KW osteoporosis; osteopathic; ss.
 XX Homo sapiens.
 XX WO200250301-A2.
 XX 27-JUN-2002.
 XX 18-DEC-2001; 2001WO-US48276.
 XX 18-DEC-2000; 2000US-255882P.
 PR 24-APR-2001; 2001US-285691P.
 XX (GENE-) GENE LOGIC INC.
 PA (PROC) PROCTER & GAMBLE CO.
 XX Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
 PI Mertz L;
 XX WPI; 2002-557663/59.
 DR Use of genes and their expression profiles associated with osteoblast
 PT differentiation for screening modulators bone formation, for diagnosing
 PT or treating e.g. osteoporosis, or as markers for the differentiation
 PT process -
 XX Claim 1; SEQ ID NO 86; 78pp + Sequence Listing; English.
 PS The invention relates to genes and their expression profiles are used
 CC for:
 CC (a) screening modulators of precursor stem cell differentiation into
 CC osteoblasts, or bone tissue deposition;
 CC (b) diagnosing abnormal deposition of bone tissue, abnormal rate of
 CC osteoblast formation or osteoporosis; or
 CC (c) treating or monitoring treatment of the conditions cited in (b), or
 CC monitoring the progression of bone tissue deposition.
 CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
 CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy,
 CC drug-induced abnormalities in bone formation or bone loss, conditions
 CC that involve altered bone metabolism (e.g. idiopathic juvenile
 CC osteoporosis), skeletal disease linked to breast cancer, mastocytosis,
 CC Fanconi syndrome or fibrous dysplasia. The present sequence is that of an
 CC osteoblast differentiation associated cDNA marker of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 160771 BP; 44776 A; 35622 C; 34804 G; 45569 T; 0 other;
 XX
 Query Match 14.8%; Score 30.4; DB 24; Length 160771;
 Best Local Similarity 54.5%; Pred. No. 26;
 Matches 61; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 47 AGCTCTGTCTTACTAAGCCAGTTCTTGTCTATTATCTTGTATCGTGGCTATGTT 106
 DB 63709 AGACCTGTCTCTACAAATAAATAATAGATGAGATTTTGAGGTTACGTTGAGCTATGAT 63768
 QY 107 AAACCCACCTAACCGTTGTTTGAAGCAGCACTGATTCCTCCAGGTATCAAAAAG 158
 DB 63769 CAGGCCACTGCGCTTTCAGCCTGGGTGACAGACGACCCCTATCTCAAAAAG 63820
 RESULT 49
 AAT58840
 ID AAT58840 standard; DNA; 580073 BP.
 XX
 AC AAT58840;
 XX
 DT 27-MAR-1997 (first entry)
 XX

DE Mycoplasma genitalium genome.
 XX
 KW M. genitalium; DNAA; DNA gyrase; origin of replication;
 XX megabase shotgun sequencing method; open reading frame; ORF; ss.
 OS Mycoplasma genitalium.
 XX
 FH Location/Qualifiers
 CDS 8552..9184
 FT /*tag= a
 FT /label= MG006
 FT /note= "Previously identified as MORF-20076, the
 FT encoded protein shows 27.59 percentage
 FT identity to thymidylate kinase (CDC8)
 FT from Saccharomyces cerevisiae"
 CDS 11252..12040
 FT /*tag= b
 FT /label= MG009
 FT /note= "Previously identified as MORF-20078, the
 FT encoded protein shows 35.43 percentage
 FT identity to the Bacillus subtilis hypothetical
 FT protein covered in accession number
 FT GB:D26185_102"
 CDS 12069..12725
 FT /*tag= c
 FT /label= MG010
 FT /note= "Previously identified as MORF-20079, the
 FT encoded protein shows 25.73 percentage
 FT identity to DNA primase (dnaE) from
 FT Clostridium acetobutylicum"
 CDS complement (13570..14247)
 FT /*tag= d
 FT /label= MG012
 FT /note= "Previously identified as MORF-20080, the
 FT encoded protein shows 31.50 percentage
 FT identity to the ribosomal protein S6
 FT modification protein (rimK) from Escherichia
 FT coli"
 CDS complement (14396..15217)
 FT /*tag= e
 FT /label= MG013
 FT /note= "Previously identified as MORF-19823, MORF-20080
 FT and MORF-20081, the encoded protein shows 33.04
 FT percentage identity to 5,10-methylene-tetra-
 FT hydrofolate dehydrogenase (fold) from E. coli"
 CDS 17474..19243
 FT /*tag= f
 FT /label= MG015
 FT /note= "Previously identified as MORF-20084, the
 FT encoded protein shows 32.23 percentage
 FT identity to transport ATP-binding protein
 FT (msbA) from E. coli"
 CDS 26478..27344
 FT /*tag= g
 FT /label= MG023
 FT /note= "Previously identified as MORF-20092, the
 FT encoded protein shows 45.96 percentage
 FT identity to fructose-bisphosphate aldolase
 FT (tsr) from B. subtilis"
 CDS 27345..28448
 FT /*tag= h
 FT /label= MG024
 FT /note= "Previously identified as MORF-19826 and
 FT MORF-20093, the encoded protein shows 46.84
 FT percentage identity to GTP-binding protein
 FT from E. coli"
 CDS 36987..38978
 FT /*tag= i
 FT /label= MG032
 FT /note= "Previously identified as MORF-20099, the
 FT encoded protein shows 26.82 percentage
 FT identity to AIP-dependent nuclease (adda)
 FT from B. subtilis"
 FT

```

FT CDS 39242..39904
FT /tag= j
FT /label= MG033
FT /note= "Previously identified as MORF-20100, the
FT encoded protein shows 35.90 percentage
FT identity to glycerol uptake facilitator
FT (glpF) from B. subtilis"
FT
FT CDS complement (39873..40514)
FT /tag= k
FT /label= MG034
FT /note= "Previously identified as MORF-20101, the
FT encoded protein shows 48.13 percentage
FT identity to thymidylate kinase (tdk)
FT from B. subtilis"
FT
FT CDS 40543..41787
FT /tag= l
FT /label= MG035
FT /note= "Previously identified as MORF-20102, the
FT encoded protein shows 30.71 percentage
FT identity to histidyl-tRNA synthetase (hiss)
FT from Mycobacterium leprae"
FT
FT CDS complement (44751..46277)
FT /tag= m
FT /label= MG038
FT /note= "Previously identified as MORF-20105, the
FT encoded protein shows 46.83 percentage
FT identity to glycerol kinase (glpk)
FT from E. coli"
FT
FT CDS complement (46268..47422)
FT /tag= n
FT /label= MG039
FT /note= "Previously identified as MORF-19831 and
FT MORF-20106, the encoded protein shows 43.20
FT percentage identity to glycerol-3-phosphate
FT dehydrogenase (GUT2) from S. cerevisiae"
FT
FT CDS 49377..49643
FT /tag= o
FT /label= MG041
FT /note= "The encoded protein shows 48.86 percentage
FT identity to phosphohistidinoprotein-hexose
FT phosphotransferase (ptsH) from Mycoplasma
FT capricolum"
FT
FT CDS 50060..51520
FT /tag= p
FT /label= MG042
FT /note= "Previously identified as MORF-19832 and
FT MORF-20108, the encoded protein shows 41.92
FT percentage identity to spermidine/
FT putrescine transport ATP-binding protein
FT (potA) from E. coli"
FT
FT CDS 51525..52382
FT /tag= q
FT /label= MG043
FT /note= "Previously identified as MORF-20110, the
FT encoded protein shows 26.51 percentage
FT identity to spermidine/putrescine transport
FT system permease protein (potB) from E. coli"
FT
FT CDS 52366..53220
FT /tag= r
FT /label= MG044
FT /note= "Previously identified as MORF-20111, the
FT encoded protein shows 29.45 percentage
FT identity to spermidine/putrescine transport
FT system permease protein C (potC) from E. coli"
FT
FT CDS 54658..55605
FT /tag= s
FT /label= MG046
FT /note= "Previously identified as MORF-20112, the
FT encoded protein shows 36.60 percentage
FT identity to sialoglycoprotease (gcp)
FT from Pasteurella haemolytica"
FT
FT CDS complement (56970..58310)
FT /tag= t

```

```

FT /label= MG048
FT /note= "Previously identified as MORF-19834,
FT MORF-20114 and MORF-20115, the encoded protein
FT shows 43.02 percentage identity to signal
FT recognition particle protein (ffh) from B.
FT subtilis"
FT CDS 58117..59079
FT /tag= u
FT /label= MG049
FT /note= "Previously identified as MORF-20114 and
FT MORF-20115, the encoded protein shows 44.78
FT percentage identity to purine-nucleoside
FT phosphorylase (deoD) from E. coli"
FT CDS 59083..59754
FT /tag= v
FT /label= MG050
FT /note= "Previously identified as MORF-20117, the
FT encoded protein shows 83.03 percentage
FT identity to deoxyribose-phosphate aldolase
FT (deoC) from Mycoplasma pneumoniae"
FT CDS complement (64898..65731)
FT /tag= w
FT /label= MG056
FT /note= "Previously identified as MORF-20122, the
FT encoded protein shows 30.25 percent
FT identity to the protein disclosed in
FT GB:D26185_99 from B. subtilis"
FT CDS complement (65713..66249)
FT /tag= x
FT /label= MG057
FT /note= "Previously identified as MORF-20123, the
FT encoded protein shows 38.90 percentage
FT identity to the protein disclosed in
FT GB:D26185_104 from B. subtilis"
FT CDS 81047..82597
FT /tag= y
FT /label= MG067
FT /note= "Previously identified as MORF-19845, the
FT encoded protein shows 28.84 percentage
FT identity to glutamic acid specific protease
FT (SPase) from Staphylococcus aureus"
FT CDS 91065..91919
FT /tag= z
FT /label= MG070
FT /note= "Previously identified as MORF-20136, the
FT encoded protein shows 34.8 percentage
FT identity to ribosomal protein S2 (rps2)
FT from Spirulina plantensis"
FT
FT Query Match 14.8%; Score 30.4; DB 18; Length 580073;
FT Best Local Similarity 53.3%; Pred. No. 41;
FT Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 37 AACAGTTAAAGTCTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGAATCGTT 96
DB 276182 AACTAGTTAAGAACTGAAATAGAAACCAATGTTCTCTCTTTTATTTGTTATTTT 276241
QY 97 GCCTATGTTAAACCCACCTAACCGTTGTTGAAGACACTGATTCCTCAGGTATCAAAA 156
DB 276242 CATCTATCTCCAAACACCCACCGTTTATTGAAAAACTGGGTTTTCACACAAAAGAA 276301
RESULT 50
AAC07910
ID AAC07910 standard; cDNA; 450 BP.
XX
AC AAC07910;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 11985.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

```

KW gene therapy; chromosome mapping; ss.
 XX Homo sapiens.
 OS EPI033401-A2.
 PN 06-SEP-2000.
 XX 21-FEB-2000; 2000EP-0200610.
 PF 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 PA Dumas Milne Edwards J, Duclert A, Giordano J;
 PI WPI; 2000-500381/45.
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 1; SEQ ID 11985; 71pp + CD-ROM; English.
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX Sequence 450 BP; 114 A; 89 C; 103 G; 140 T; 4 other;
 SQ

Query Match 14.7%; Score 30.2; DB 21; Length 450;
 Best Local Similarity 50.3%; Pred. No. 3.4;
 Matches 74; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 QY 35 AGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCTTGATTCTG 94
 Db 147 AGAAAGGGTTAAAGGCTGATGGTACCTAAGGCTGGTACTTGAATTTTCATCAAGATAAG 206
 QY 95 TTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTTGCCAGGTATCAA 154
 Db 207 CTGCGCTTAAGTTCTCTTCATTACACAAATGATCCTAGATAATTGATAGATCCTGTGGTTC 266
 QY 155 AAGTCTGTGAAGGTTCTCTCGGGTAT 181
 Db 267 AACTGGATTCTAGATAGAGCTGGAT 293

RESULT 51
 AAH64937
 ID AAH64937 standard; cDNA; 1489 BP.
 XX AAH64937;
 AC
 XX 11-SEP-2001 (first entry)
 DT
 XX Human secreted protein cDNA, SEQ ID NO: 213.
 DE
 XX Human, secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KW GENSET; ss.
 XX Homo sapiens.
 OS

PN WO200142451-A2.
 XX 14-JUN-2001.
 XX 07-DEC-2000; 2000WO-IB01938.
 PF 08-DEC-1999; 99US-0169629.
 PR 06-MAR-2000; 2000US-0187470.
 XX (GEST) GENSET.
 PA Dumas Milne Edwards J, Bougueleret L, Jobert S;
 PI WPI; 2001-367870/38.
 XX P-PSDB; AAG89334.
 DR Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX Claim 7; Page 766-767; 921pp; English.
 XX The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patient's own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET nucleic acid of the invention.
 XX
 SQ Sequence 1489 BP; 425 A; 323 C; 335 G; 406 T; 0 other;
 Query Match 14.7%; Score 30.2; DB 22; Length 1489;
 Best Local Similarity 50.3%; Pred. No. 5.3;
 Matches 74; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 QY 35 AGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCTTGATTCTG 94
 Db 145 AGAAAGGGTTAAAGGCTGATGGTACCTAAGGCTGGTACTTGAATTTTCATCAAGATAAG 204
 QY 95 TTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTTGCCAGGTATCAA 154
 Db 205 CTGCGCTTAAGTTCTCTTCATTACACAAATGATCCTAGATAATTGATAGATCCTGTGGTTC 264
 QY 155 AAGTCTGTGAAGGTTCTCTCGGGTAT 181
 Db 265 AACTGGATTCTAGATAGAGCTGGAT 291

RESULT 52
 ABN85329/c
 ID ABN85329 standard; DNA; 36159 BP.
 XX ABN85329;
 AC
 XX 04-OCT-2002 (first entry)
 DT
 XX Human kinase gene.
 DE
 XX Human, kinase; enzyme; homeodomain-interacting protein kinase; testis;
 KW brain medulloblastoma; infant brain; schizophrenia; retina; gene;
 KW germinal center B cell; colon; liver; chromosome 1; gene therapy;
 KW

KW SNP; single nucleotide polymorphism; ds.
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT CDS 2490..33161
 FT /tag= a
 FT /product= "Human kinase"
 FT /note= "Contains 14 introns"
 FT 2490..3565
 FT /tag= b
 FT /number= 1
 FT /cons_splice= (5'site:YES,3'site:NO)
 FT 3566..14639
 FT /tag= c
 FT /number= 1
 FT replace(4452,T)
 FT /tag= d
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(5330,G)
 FT /tag= e
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(9256,A)
 FT /tag= f
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(11773,G)
 FT /tag= g
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(12886,G)
 FT /tag= h
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(14130..14131,A-)
 FT /tag= i
 FT /standard_name= "Single nucleotide polymorphism"
 FT 14640..14763
 FT /tag= j
 FT /number= 2
 FT 14764..17436
 FT /tag= k
 FT /number= 2
 FT 17437..17523
 FT /tag= l
 FT /number= 3
 FT 17524..18501
 FT /tag= m
 FT /number= 3
 FT 18502..18686
 FT /tag= n
 FT /number= 4
 FT 18687..18998
 FT /tag= o
 FT /number= 4
 FT 18999..19161
 FT /tag= p
 FT /number= 5
 FT /transl_except= (pos:19063..19065,aa:Asn)
 FT 19162..19940
 FT /tag= q
 FT /number= 5
 FT 19941..20166
 FT /tag= r
 FT /number= 6
 FT 20167..25525
 FT /tag= s
 FT /number= 6
 FT 25526..25647
 FT /tag= t
 FT /number= 7
 FT 25648..26595
 FT /tag= u
 FT /number= 7
 FT 26596..26730
 FT /tag= v
 FT
 FT intron /number= 8
 FT 26731..29338
 FT /tag= w
 FT /number= 8
 FT 29339..29482
 FT /tag= x
 FT /number= 9
 FT /cons_splice= (5'site:YES,3'site:NO)
 FT /note= "The location for this exon was given as
 FT 29339-29487 in the specification"
 FT 29483..30975
 FT /tag= y
 FT /number= 9
 FT 30976..31157
 FT /tag= z
 FT /number= 10
 FT /cons_splice= (5'site:NO,3'site:YES)
 FT /note= "The location for this exon was given as
 FT 30975-31157 in the specification"
 FT 31158..31650
 FT /tag= aa
 FT /number= 10
 FT 31651..31857
 FT /tag= ab
 FT /number= 11
 FT 31858..32003
 FT /tag= ac
 FT /number= 11
 FT 32004..32171
 FT /tag= ad
 FT /number= 12
 FT /note= "32004-32171 is the location given in the
 FT specification, but this location does not concur with
 FT the protein encoded by this sequence"
 FT 32172..32402
 FT /tag= ae
 FT /number= 12
 FT 32403..32436
 FT /tag= af
 FT /number= 13
 FT /note= "32403-32436 is the location given in the
 FT specification, but this location does not concur with
 FT the protein encoded by this sequence"
 FT 32437..32440
 FT /tag= ag
 FT /number= 13
 FT 32441..32617
 FT /tag= ah
 FT /number= 14
 FT /note= "32441-32617 is the location given in the
 FT specification, but this location does not concur with
 FT the protein encoded by this sequence"
 FT 32618..32672
 FT /tag= ai
 FT /number= 14
 FT 32673..33161
 FT /tag= aj
 FT /number= 15
 FT
 FT intron WO200253717-A2.
 FT 11-JUL-2002.
 FT 19-DEC-2001; 2001WO-US48534.
 FT 28-DEC-2000; 2000US-0749588.
 FT (PEKE) PE CORP NY.
 FT Chandramouliswaran I, Guegler K, Webster M, Yan C, Di Francesco V;
 FT Beasley EM;
 FT WPI; 2002-583610/62.
 FT DR

DR P-PSDB; ABB83490.

XX Human kinase protein, related to homeodomain-interacting protein kinase

PT subfamily, useful as a model for developing human therapeutic targets

PT and serves as a target for human therapeutics

XX Claim 22; Fig 3; 96pp; English.

XX The present sequence is the gene for a human kinase, which is related to

CC the homeodomain-interacting protein kinase subfamily. The kinase and its

CC DNA sequence can be used as models for the development of human

CC therapeutic targets, aid in the identification of therapeutic proteins

CC and serve as targets for the development of human therapeutic agents that

CC modulate kinase activity in cells and tissues that express the kinase.

CC Experimental data indicates expression of the human kinase in testis,

CC brain medulloblastomas, infant brain, schizophrenic brain, retina,

CC germinal center B cells, colon and liver. The human kinase gene is

CC located on chromosome 1.

XX Sequence 36159 BP; 8425 A; 6177 C; 6348 G; 10756 T; 4453 other;

Query Match 14.7%; Score 30.2; DB 24; Length 36159;

Best Local Similarity 49.1%; Pred. No. 17;

Matches 80; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 31 CTCAGAACCCAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGA 90

DB 12268 CTCAGATCCAGATCCAGATCCAGATTCAGATTCCTCATGTTACAGGAGTTAAATGGAAA 12209

QY 91 TTCGTTGGCCTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATGTCAGGTA 150

DB 12208 ACAGATCTTTCGCTGAAGAACCTACTAAGTGGCCGCTGCTCTCACAGATACCCAGTT 12149

QY 151 TCAGAAAGTCTGTGAGGTTCTCGGTATGCGTTGCTTTTCTGT 193

DB 12148 TCTCAAACTACTTCAAGGATATGCTTATTGTTGTTTGT 12106

RESULT 53

AAK53717

ID AAK53717 standard; cDNA; 344 BP.

XX AAK53717;

XX 16-NOV-2001 (first entry)

DE Murine replication associated protein encoding cDNA SEQ ID 282.

XX Murine; liver; gene library; amino acid synthesis; binding protein;

KW cell metabolism; energy metabolism; fatty acid metabolism; synthesis;

KW phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;

KW replication; transcription; translation; transport protein; ss.

XX Mus musculus.

XX DE20103510-U1.

XX 07-JUN-2001.

XX 28-FEB-2001; 2001DE-2003510.

XX 02-DEC-1999; 99DE-1058160.

XX (LION-) LION BIOSCIENCE AG.

XX WPI; 2001-368570/39.

PT Gene library containing sequences with specific 3'-ends and no polyA

PT tail, encoding proteins involved in a wide range of cellular processes

XX 'Claim 15; Page 97; 251pp; German.

CC This invention describes a novel gene library (A) comprises a gene

CC sequence (or its part) encoding a protein involved in amino acid

CC synthesis, cellular/energy metabolism, metabolism of

CC fatty acids/phospholipids, synthesis or breakdown of

CC purines/pyrimidines/nucleosides/nucleotides, DNA

CC replication/transcription/translation, or is a transport/binding protein.

CC (A) are produced that correspond to the 3'-end of mRNA but without the

CC polyA tail. They can be prepared more efficiently and with less effort

CC than conventional libraries. AAK53436-AAK54275 represent fragments of the

CC gene library described in the method of the invention.

XX Sequence 344 BP; 73 A; 78 C; 70 G; 123 T; 0 other;

Query Match 14.5%; Score 29.8; DB 22; Length 344;

Best Local Similarity 56.7%; Pred. No. 4.2;

Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 82 TTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGACACTGATT 141

DB 230 TTAGCTGGTCTTTTCCGCCATGTTTATACACCATATATGCTTCGTCGCCCGGCTG 289

QY 142 GTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178

DB 290 ATCTGATAGCAGAGATTCCTCTCTGCACCCGAGG 326

RESULT 54

AAT68952/c

ID AAT68952 standard; DNA; 5150 BP.

XX AAT68952;

XX 05-AUG-1997 (first entry)

XX Blackcurrant fruit-specific RIB7 gene promoter.

XX Blackcurrant; fruit-specific promoter; RIB7 gene; transgenic plant;

XX ss.

XX Ribes nigrum strain Ben Alder.

XX Key Location/Qualifiers

FT TATA_signal 3041..3044

FT /*tag= a

FT CDS 3156

FT /*tag= b

FT /codon_start= 3156

FT /note= "putative ATG translational start codon"

XX WO9717452-A1.

XX 15-MAY-1997.

XX 04-NOV-1996; 96WO-EP04807.

XX 03-NOV-1995; 95GB-0022558.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Brennan RM, Taylor MA, Woodhead MR;

XX WPI; 1997-281041/25.

XX New isolated promoters from blackcurrant fruit - used for driving

PT fruit-specific expression of DNA sequences in transgenic

PT blackcurrant and other non-climacteric fruit

XX Claim 3; Page 44-48; 66pp; English.

XX A DNA sequence (AAT68952) comprises the promoter region of the

CC RIB7 gene. This gene was isolated from a blackcurrant genomic

CC library using a RIB7 cDNA clone (see also AAT68957) as probe. The

CC RIB7 gene is highly expressed in ripe blackcurrant fruit and

CC expressed at negligible levels in other tissues of the blackcurrant
 CC plant. Its promoter region can be used to drive fruit-specific
 CC expression of cloned downstream DNA sequences in transgenic
 CC blackcurrant or other non-climacteric fruit as a means of
 CC manipulating the ripening process.
 XX
 SQ Sequence 5150 BP; 1599 A; 1069 C; 899 G; 1583 T; 0 other;
 Query Match 14.3%; Score 29.4; DB 18; Length 5150;
 Best Local Similarity 55.3%; Pred. No. 16;
 Matches 57; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 100 CTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTCAGGTATCAAAAAGT 159
 Db 2569 CTATTTAAACAAGACTTCTAAGTGTGTTGAGGATATTATTCATATTTATCAATAAAAA 2510
 QY 160 GCTGTGAAGGTTCTCTGCGGTATGCTGTGTTTCGTTCCACAATA 202
 Db 2509 ACTGAGAACTTTTCTCGTGTCTCTCTCTCTTCTTACGAGNAYA 2467
 RESULT 55
 ABQ89443
 ID ABQ89443 standard; cDNA; 756 BP.
 XX
 AC ABQ89443;
 XX
 DT 27-SEP-2002 (first entry)
 XX
 DE Human prostate expressed polynucleotide SEQ ID NO 699.
 XX
 DE Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy;
 KW Gene; ss.
 KW
 XX Homo sapiens.
 OS
 XX WO20025700-A2.
 PN
 XX 18-JUL-2002.
 PD
 XX
 XX 07-DEC-2001; 2001WO-US47349.
 XX
 PR 07-DEC-2000; 2000US-254648P.
 PR 13-MAR-2001; 2001US-275688P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
 PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones WL, Stache-Crain B, Scott EM;
 XX
 DR WPI; 2002-557824/59.
 XX
 XX New genes and gene products isolated from human prostate, useful for
 PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
 PT cancer), or as vaccines for treating or preventing these diseases -
 XX
 PS Claim 1; SEQ ID NO 699; 186pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide comprising any of
 CC 1477 sequences or its fragment, degenerate variant, antisense or
 CC complement. The polynucleotides and gene products are useful for
 CC or diagnosing tumor or cancer (e.g. prostate cancer, breast cancer,
 CC lung cancer or medullary carcinoma) in a subject (e.g. cattle, dogs,
 CC cats, rabbits, horse or human). The polynucleotides and polypeptides are
 CC also useful as vaccines for treating or preventing these diseases. The
 CC polynucleotides are useful for gene therapy. The present sequence is that
 CC of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
 CC electronically as sequences of the invention. However only 1271
 CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
 CC proteins are claimed.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequence.
 XX
 SQ Sequence 756 BP; 236 A; 138 C; 158 G; 214 T; 10 other;
 Query Match 14.2%; Score 29.2; DB 24; Length 756;
 Best Local Similarity 49.4%; Pred. No. 9;
 Matches 76; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 QY 38 ACCAGTTAAAGTCTCTGTCTACTAAGCCAGGTTCTTGCTCTATTATCTTGATTCGTTG 97
 Db 291 ACTATTTAAATCATCTGCGCTTACTACAAACCTTTTCTGCACTGACTTCTTCTGTG 350
 QY 98 CGCTATGTTAAACCCACCTAACCGTGTGTTTGAAGGACACTGATTCAGGTATCAAAAA 157
 Db 351 TGTGATTTCAAAATTAATTTTAATGTTTACTAAAAACAAGTTATCTACCATTTATTATAA 410
 QY 158 GTGCTGTGAAGGTTCTGCGGTATGCGTGTGTTTC 191
 Db 411 ACTCAATTGTAGGCCAGGCGTGTGCTTACGTC 444
 RESULT 56
 AAS00193/C
 ID AAS00193 standard; DNA; 549 BP.
 XX
 AC AAS00193;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE S. aureus DNA encoding transcription antitermination factor, NusG.
 XX
 DE Transcription antitermination factor; NusG; immunogen; vaccine;
 KW antibody; wound infection; cellulitis; burn infection; eyelid infection;
 KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
 KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
 KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
 ds.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..549
 FT /*tag= a
 FT /product= "NusG"
 FT /partial
 FT /note= "No stop codon"
 XX
 XX WO200116292-A2.
 XX
 XX 08-MAR-2001.
 XX
 XX 31-AUG-2000; 2000WO-US23773.
 XX
 PR 01-SEP-1999; 99US-0151933.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Choi GH;
 XX
 DR WPI; 2001-183259/18.
 DR P-PSDB; AAU00832.
 XX
 XX New isolated nucleic acid for use in diagnosing Staphylococcus
 PT infections and in vaccines for eliciting immune responses to the
 PT infections -
 XX
 PS Claim 1; Page 15; 225pp; English.
 XX
 CC The sequence encodes S. aureus NusG (transcription antitermination
 CC factor). The polynucleotides of the invention are used to detect
 CC Staphylococcus nucleic acids in a biological sample from an animal for
 CC diagnosing Staphylococcus infections. The polypeptides of the invention

CC are used to detect anti-Staphylococcus antibodies in a biological sample
 CC from an animal to diagnose Staphylococcus infections. The polypeptides
 CC are also used in vaccines to elicit protective antibodies in an animal to
 CC a member of the Staphylococcus genus and for preventing or attenuating an
 CC infection caused by a member of the Staphylococcus genus e.g wound
 CC infection, cellulitis, burn infection, eyelid infection, food poisoning,
 CC joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
 CC scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
 CC disease and Lyell's disease), toxic shock syndrome and endocarditis.
 CC The polynucleotides may also be used in vaccines and for preventing or
 CC attenuating a Staphylococcus infection. Antibodies to the polypeptides
 CC may be used to purify, detect and target the polypeptides in vitro and
 CC in vivo diagnostic and therapeutic methods.

XX Sequence 549 BP; 206 A; 73 C; 125 G; 145 T; 0 other;

Query Match 14.1%; Score 29; DB 22; Length 549;
 Best Local Similarity 51.1%; Pred. No. 9.3;
 Matches 66; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 2 ATTCGAGTCGGTACCATACCTGCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTAC 61
 |||||
 Db 315 ATTTGGCTTAGACCTGCACCTGCAGAACCTACAAACCCAGTAACGCTGGTGTATTCT 256
 |||||
 QY 62 TAAGCCAGGTTCTTGTCTATTATCTTGAATTCGTTGGCTATGTTAAACCCACCTAACCG 121
 |||||
 Db 255 TACCACATACCATTGATTCATCTGTCATGATTAATTCCACTAAACATATCCAGGGAATGT 196
 |||||
 QY 122 TTGTTTGAAGGAC 134
 |||||
 Db 195 TTTTAAACAGTC 183
 |||||

RESULT 57

AAS55051/c

ID AAS55051 standard; DNA; 549 BP.

AC AAS55051;

DT 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #1363.

XX Antisense; ds; prokaryotic cellular proliferation gene;
 XX antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI, 2001-611495/70.

DR P-PSDB; AAU37192.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -

PS Claim 27; Seq ID No 8688; Slipp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 549 BP; 208 A; 73 C; 122 G; 146 T; 0 other;

Query Match 14.1%; Score 29; DB 23; Length 549;
 Best Local Similarity 51.1%; Pred. No. 9.3;
 Matches 68; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 2 ATTCGAGTCGGTACCATACCTGCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTAC 61
 |||||
 Db 312 ATTTGGCTTAGACCTGCACCTGCAGAACCTACAAACCCAGTAACGCTGGTGTATTCT 253
 |||||
 QY 62 TAAGCCAGGTTCTTGTCTATTATCTTGAATTCGTTGGCTATGTTAAACCCACCTAACCG 121
 |||||
 Db 252 TACCACATACCATTGATTCATCTGTCATGATTAATTCCACTAAACATATCCAGGGAATGT 193
 |||||
 QY 122 TTGTTTGAAGGAC 134
 |||||
 Db 192 TTTTAAACAGTC 180
 |||||

RESULT 58

AAS55319/c

ID AAS55319 standard; DNA; 549 BP.

XX AAS55319;

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #1631.

XX Antisense; ds; prokaryotic cellular proliferation gene;
 XX antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI

PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241807.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0255678.
 PR XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -

PS Disclosure; SEQ ID NO 23303; 3071pp + Sequence Listing; English.

XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 CC XX

SQ Sequence 32145 BP; 8777 A; 7496 C; 7117 G; 8755 T; 0 other;

Query Match 14.1%; Score 29; DB 22; Length 32145;

Best Local Similarity 49.7%; Pred. No. 42;

Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 23 TGCATATGCTCAAGAACCCAGTTAAAGGTCCTGCTGCTACTAAGCCAGGTTCTTGTCTTAT 82

Db 9310 TTCCCTCGCTGTTGGCCATGCTATGTCCTTTTGAGAAGTGTCTGTCATGTCCTTT 9369

QY 83 TATCTTGATTGCTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTG 142

Db 9370 GCTCATTTTTTAATGGGGTGTGTTAAATTGCAGATATTAGACCTTAGATGATGCCAGAGTT 9429

QY 143 TCCAGGTATCAAAAAGTCTGTGAAGGTT 171

Db 9430 TCGAAATATTTTCTCATTTTGTAGGTT 9458

RESULT 60

AAK68575

ID AAK68575 standard; DNA; 32145 BP.

AC AAK68575;

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23387.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 32145 BP; 8777 A; 7496 C; 7117 G; 8755 T; 0 other;

Query Match 14.1%; Score 29; DB 22; Length 32145;
 Best Local Similarity 49.7%; Pred. No. 42;
 Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 23 TGCATATGCTCAAGAACACAGTTAAAGTCTGCTCTACTAAGCCAGGTTCTTGTCCTAT 82
 Db 9310 TTCTCTCTGTTGGCCATGATGTCCTTTCTTTTGGAGAAGTGTCTGTTTCATGTCCTTT 9369

QY 83 TATCTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTG 142
 Db 9370 GCTCATTTTTTAATGGGGTCTGTTTAAATGTCAGATATAGACCTTAGATGATCCAGATT 9429

QY 143 TCCAGGTATCAAAAAGTCTGTGAAGTT 171
 Db 9430 TGCAATATTTTGTCTCATTTTGTAGTT 9458

RESULT 61
 ID ABA90521/c
 ID ABA90521 standard; DNA; 2365589 BP.
 AC ABA90521;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Genomic sequence of *Lactococcus lactis* IL1403.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.
 XX
 OS *Lactococcus lactis* IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX
 DR WPI; 2002-043418/06.
 XX
 PT New nucleotide sequence useful in the identification or *Lactococcus*
 PT *lactis* and related species -
 XX
 PS Claim 1; SEQ ID 1; 2504pp; French.
 XX
 CC The present invention is related to a *Lactococcus lactis* nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify *Lactococcus lactis* or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;

Query Match 14.1%; Score 29; DB 24; Length 2365589;

Best Local Similarity 51.1%; Pred. No. 1.9e+02;
 Matches 68; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 34 AAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCTTGATTC 93
 Db 2287194 AAGAGGTAAATATGAAATGAATTTGAAAAACAAATCTTCTCTTATCAACTTTACTT 2287135

QY 94 GTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCAGGTATCA 153
 Db 2287134 GTTACGACAGGTGCGATAGGGGTAACCTTCTGTATCTCGCACTTAAGTCGCTATGAAAA 2287075

QY 154 AAAAGTGTGTGA 166
 Db 2287074 ACAAGTTCAGTTA 2287062

RESULT 62
 AAC56088
 ID AAC56088 standard; DNA; 1677 BP.
 XX
 AC AAC56088;
 XX
 DT 25-JAN-2001 (first entry)
 XX
 DE *Eucalyptus grandis* transcription factor DNA sequence #219.
 XX
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB; ss.
 XX
 OS *Eucalyptus grandis*.
 XX
 PN WO200053724-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06112.
 XX
 PR 11-MAR-1999; 99US-0266513.
 PR 18-AUG-1999; 99US-0149485.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Wood M, McGrath A, Shenk MA, Glenn M;
 XX
 DR WPI; 2000-579369/54.
 XX
 PT New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT -
 XX
 PS Claim 1; Pages 103-104; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC *Eucalyptus grandis* or *Pinus radiata*. The present sequence is the coding
 CC sequence for one such transcription factor. The transcription factor may
 CC be used to produce a plant having modified gene expression such as a
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
 CC mahogany species or to modify the activity of a polypeptide in a plant.
 CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bZIP, bZIP family of G-box
 CC binding factors, basic helix-loop-helix zipper,
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
 CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
 CC and MYB.
 XX
 SQ Sequence 1677 BP; 424 A; 371 C; 433 G; 449 T; 0 other;

Query Match 14.0%; Score 28.8; DB 21; Length 1677;

DE Human bone marrow expressed single exon probe SEQ ID NO: 2212.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 2212; 659pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
 Query Match 13.8%; Score 28.4; DB 22; Length 464;
 Best Local Similarity 54.9%; Pred. No. 14;
 Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 77 TCCTATTATCTGATCGTGGCTATGTTAAACCCACCTAACCGTTCTTGAAGGACAC 136
 DB 293 TCCTCTTTCCACAGATTATTGCTGCTCTCTCCAGCCCTTCACTGTAGATGGAAGCCTG 352
 QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
 DB 353 TGATTTTCTGGCATCATGCTGACTGAGAGGTCAGTGTGG 394
 QY 17 OCT-2001 (first entry)
 DE Probe #2168 for gene expression analysis in human cervical cell sample.
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 25; SEQ ID No 2168; 487pp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human HeLa cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
 Query Match 13.8%; Score 28.4; DB 22; Length 464;
 Best Local Similarity 54.9%; Pred. No. 14;
 Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 77 TCCTATTATCTGATCGTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACAC 136
 DB 293 TCCTCTTTCCACAGATTATTGCTGCTCTCTCCAGCCCTTCACTGTAGATGGAAGCCTG 352
 QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
 DB 353 TGATTTTCTGGCATCATGCTGACTGAGAGGTCAGTGTGG 394
 RESULT 69
 AA133589
 ID AA133589 standard; DNA; 464 BP.
 XX
 AC AA133589;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #2275 used to measure gene expression in human placenta sample.
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR

PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 PT Claim 25; SEQ ID No 2275; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
 Query Match 13.8%; Score 28.4; DB 22; Length 464;
 Best Local Similarity 54.9%; Pred. No. 14;
 Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 77 TCCTATTATCTTGATTCGTTGGCTATGTTAAACCCACTTAACCGTTGTTTGAAGGACAC 136
 Db 293 TCCTCTTTCCACAGATTATTGCTGCTCTCTCAAGCCCTTCACTGTAGATGAAGCCTG 352
 QY 137 TGATTCTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
 Db 353 TGATTTTCTGGCAATCATGCTACTGAGAGGTCAGTGTGG 394
 RESULT 70
 AA102148
 ID AA102148 standard; DNA; 464 BP.
 AC AA102148;
 XX 09-OCT-2001 (first entry)
 DT Probe #2139 used to measure gene expression in human breast sample.
 XX Probe; human; breast disease; breast cancer; development disorder; ss;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX Homo sapiens.
 OS WO200157270-A2.
 PN 09-AUG-2001.
 PD 29-JAN-2001; 2001WO-US00661.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.
 DR Novel single exon nucleic acid probe used to measuring gene expression
 PT 'in a human breast -
 XX

PS Claim 25; SEQ ID No 2139; 322pp; English.
 XX The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
 Query Match 13.8%; Score 28.4; DB 22; Length 464;
 Best Local Similarity 54.9%; Pred. No. 14;
 Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 77 TCCTATTATCTTGATTCGTTGGCTATGTTAAACCCACTTAACCGTTGTTTGAAGGACAC 136
 Db 293 TCCTCTTTCCACAGATTATTGCTGCTCTCTCAAGCCCTTCACTGTAGATGAAGCCTG 352
 QY 137 TGATTCTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
 Db 353 TGATTTTCTGGCAATCATGCTACTGAGAGGTCAGTGTGG 394
 RESULT 71
 ABS02116
 ID ABS02116 standard; DNA; 464 BP.
 XX ABS02116;
 XX 19-AUG-2002 (first entry)
 DT Human genome-derived single exon probe from lung SEQ ID No 2107.
 DE Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS WO200186003-A2.
 PN 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 DR

XX OS Homo sapiens.
 XX KW WO200177290-A2.
 XX PN 18-OCT-2001.
 XX PD
 XX PF 29-MAR-2001; 2001WO-US10295.
 XX PR 06-APR-2000; 2000US-194941P.
 XX PA (GEMY) GENETICS INST INC.
 XX PI Wong GG, Clark HF, Fichtel K, Agostino MJ, Howes SH, Resnick RJ;
 XX PI Gulukota K, Graham JR;
 XX DR WPI; 2002-179323/23.
 XX
 XX PT Six hundred and twenty five polynucleotides derived from a variety of
 XX PT human tissue sources which encode secreted proteins, useful for
 XX PT treating immune deficiencies and disorders such as autoimmune disorders
 XX PT
 XX PS Claim 1; Page 273; 339pp; English.
 XX
 XX CC The invention relates to 625 polynucleotides which have been derived from
 XX CC a variety of human tissue sources and which encode novel secreted
 XX CC proteins, their complements and sequences that hybridise to them.
 XX CC Also included are a vector comprising the polynucleotide, a host cell
 XX CC transformed with the vector, the proteins encoded by the
 XX CC polynucleotides, antibodies that bind to the proteins and identification
 XX CC of modulators of the proteins or the expression of the polynucleotide.
 XX CC The polynucleotides can be used as probes for the identification
 XX CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 XX CC and proteins can also be used as nutritional supplements. The protein
 XX CC is useful in the treatment of various immune deficiencies and disorders
 XX CC such as viral infections, bacterial infections, fungal infections,
 XX CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 XX CC autoimmune thyroiditis and diabetes) and allergic reactions and
 XX CC conditions (e.g. asthma). They are also useful for treating
 XX CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 XX CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 XX CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 XX CC useful for tissue regeneration, for wound healing and in the treatment
 XX CC of burns, incisions and ulcers. The proteins are also useful for
 XX CC regulating haematopoiesis, for treating myeloid or lymphoid cell
 XX CC deficiencies. The present sequence is one of the 625 cDNA sequences
 XX CC encoding a secreted protein.
 XX
 XX SQ Sequence 1636 BP; 418 A; 411 C; 351 G; 456 T; 0 other;
 Query Match 13.8%; Score 28.4; DB 24; Length 1636;
 Best Local Similarity 62.9%; Pred. No. 22;
 Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 QY 32 TCAGAACAGTTAAAGGTCCTGCTACTAAGCCAGGTTCTTCTCCTATATCTTGAT 91
 Db 1104 TCAAGCTACAGGTAAGGCTCTACCAACCAATTAAGGTAATTTGACTTATCCTATGAT 1045
 QY 92 TCGTTGGCT 101
 Db 1044 TCCTTCTTCT 1035
 RESULT 74
 AAX60810/c
 ID AAX60810 standard; DNA; 1656 BP.
 XX AC AAX60810;
 XX XX 09-AUG-1999 (first entry)
 XX DT
 XX DE Human secreted protein encoding DNA (clone yb8-1).

XX KW Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
 XX KW nutritional activity; cytokine; cell proliferation; immune stimulation;
 XX KW hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
 XX KW anti-inflammatory; tumour invasion; ss.
 XX OS Homo sapiens.
 XX PN WO9926961-A1.
 XX PD 03-JUN-1999.
 XX PF 24-NOV-1998; 98WO-US25149.
 XX PR 23-NOV-1998; 98US-0197886.
 XX PR 26-NOV-1997; 97US-0066804.
 XX PA (GEMY) GENETICS INST INC.
 XX PI Agostino MJ, Clark HF, Collins-Racie LA, Evans C;
 XX PI Fichtel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 XX PI Steininger RJ, Treacy M, Wong GG;
 XX DR WPI; 1999-357809/30.
 XX DR P-PSDB; AAY17228.
 XX PT New polynucleotides encoding secreted proteins
 XX PS Claim 39; Page 125-126; 133pp; English.
 XX CC The invention relates to secreted proteins (AAY17219-228) encoded by
 XX CC polynucleotides obtained from human fetal kidney, adult lung, adult
 XX CC kidney, adult brain, adult blood, adult testes, and fetal brain and
 XX CC murine adult bone marrow cDNA libraries. The secreted protein nucleic
 XX CC acid sequences (X6801-811) correspond to clones b0306-7, gJ283-6,
 XX CC fk317-3, k213-2x, na316-1, nf93-20, np164-1, pe204-1, yal-1 and yb-1,
 XX CC (all clones are deposited as ATCC 98599); The PNs and proteins are
 XX CC predicted to have biological activities which would make them suitable
 XX CC for treating, preventing or ameliorating medical conditions in humans and
 XX CC animals, although no supporting data is given. Suggested activities
 XX CC include nutritional activity, cytokine and cell
 XX CC proliferation/differentiation activity, immune stimulating (e.g. as
 XX CC vaccines) or suppressing activity, hematopoiesis regulating activity,
 XX CC tissue growth activity, activin/inhibin activity,
 XX CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 XX CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 XX CC invasion suppressor activity, and tumour inhibition activity. The PNs are
 XX CC also stated to be useful for gene therapy.
 XX
 XX SQ Sequence 1656 BP; 473 A; 389 C; 340 G; 454 T; 0 other;
 Query Match 13.8%; Score 28.4; DB 20; Length 1656;
 Best Local Similarity 62.9%; Pred. No. 22;
 Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 QY 32 TCAGAACAGTTAAAGGTCCTGCTACTAAGCCAGGTTCTTCTCCTATATCTTGAT 91
 Db 1023 TCAAGCTACAGGTAAGGCTCTACCAACCAATTAAGGTAATTTGACTTATCCTATGAT 964
 QY 92 TCGTTGGCT 101
 Db 963 TCCTTCTTCT 954
 RESULT 75
 AAS59216/c
 ID AAS59216 standard; cDNA; 1656 BP.
 XX AC AAS59216;
 XX XX 16-JAN-2002 (first entry)
 XX DT
 XX DE Human cDNA encoding a secreted protein yb8_1.

XX Human; secreted protein; ss; antiinflammatory; immunosuppressive;
 KW neurotropic; neuroprotective; antiarthritic; antimicrobial; vulnerary;
 KW cytoskeletal; antidiabetic; virucide; antiinfectivity; anticonvulsant;
 KW vasotrophic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; antiulcer; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KW food supplement.
 XX
 OS Homo sapiens.
 XX
 PN WO200175068-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 22-MAR-2001; 2001WO-US09369.
 XX
 PR 30-MAR-2000; 2000US-0539330.
 XX
 PR 04-DEC-2000; 2000US-0729674.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 PI Treacy M, Agostino M, Steininger RJ, Spaulding V, Wong GG;
 PI Clark H, Fechtel K, Merberg D;
 XX
 DR WPI; 2001-639363/73.
 DR P-PSDB; AAU38998.
 XX
 PT Secreted human proteins, useful as vaccine for treating various
 PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 PT nervous system disorders (e.g. stroke) -
 XX
 PS Claim 13; Page 471; 619pp; English.
 XX
 CC The invention relates to novel human secreted proteins, the nucleic
 CC acids encoding them. The protein may exhibit cytokine, cell proliferation
 CC or cell differentiation activity or may induce production of other
 CC cytokines in certain cell populations and may exhibit immune stimulating
 CC or immune suppressing activity, which is useful for the treatment of
 CC various immune deficiencies and disorders e.g. severe combined
 CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation. The proteins are also useful in the treatment of diseases
 CC and disorders including tissue, skin and organ transplantation and in
 CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
 CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
 CC in the treatment of burns, incisions and ulcers; as well as in treatment
 CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
 CC inflammatory processes, diseases of the peripheral nervous system,
 CC Alzheimer's, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, and Shy-Drager syndrome,
 CC infarction of cardiac and central nervous system vessel e.g. stroke,
 CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 CC protein, having activin- or inhibin-related activities is useful as a
 CC contraceptive based on the ability of inhibins to decrease fertility in
 CC female mammals and decrease spermatogenesis in male mammals. The
 CC proteins and nucleic acids are also useful as food supplements. The
 CC present sequence encodes a secreted protein of the invention.
 XX
 SQ Sequence 1656 BP; 473 A; 389 C; 340 G; 454 T; 0 other;

Query Match 13.8%; Score 28.4; DB 22; Length 1656;
 Best Local Similarity 62.9%; Pred. No. 22;
 Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 32 TCAAGAACCAAGTTAAAGGCTGTGTCTACTAAGCCAGGTTCTTGTCTATTATCTTGAT 91

Db 1023 TCAAGCTACAGGTAAGGCTCTACCAACATTAAAGGTATTGACTTATCTATGAATTGAT 964
 OY 92 TCGTTGGCGT 101
 Db 963 TCTTTCTTCT 954

Search completed: February 15, 2003, 22:16:54
 Job time : 811 secs

